
Zur Berechnung der Wahrscheinlichkeitsdichte gefilterter stochastischer Prozesse

vorgelegt dem Fachbereich 1 (Elektrotechnik / Physik)
der Universität Bremen als Dissertation zur Erlangung des Grades
eines Doktors der Ingenieurwissenschaften (Dr.-Ing.)

von

Klaus Uwe Simmer

1968



Zur Berechnung der Wahrscheinlichkeitsdichte gefilterter stochastischer Prozesse

vorgelegt dem Fachbereich 1 (Elektrotechnik / Physik)
der Universität Bremen als Dissertation zur Erlangung des Grades
eines Doktors der Ingenieurwissenschaften (Dr.-Ing.)

von

Klaus Uwe Simmer

Tag des öffentlichen Kolloquiums:

20. 10. 88

Gutachter der Dissertation:

Prof. Dr. A. Wasiljeff

Prof. Dr. F. Arndt

<u>Inhaltsverzeichnis</u>	<u>Seite</u>
1. Einleitung.....	1
2. Vorarbeiten.....	3
2.1 Vorarbeiten anderer Autoren.....	3
2.2 Eigene Vorarbeiten.....	13
3. Filterung stochastischer Prozesse.....	22
3.1 Rauschen und stochastische Prozesse.....	24
3.2 Lineare Filter.....	26
3.3 Charakteristische Funktion und Verteilungsfunktion...	33
3.4 Wahrscheinlichkeitsdichte nach der Filterung.....	43
3.5 Berechnung mit Hilfe der multidimensionalen Dichte...	46
3.6 Charakteristische Funktion nach der Filterung.....	50
4. Anwendung auf einen quadratischen Detektor.....	62
4.1 Charakteristische Funktion nach der ZF - Filterung..	63
4.2 Der Quadrierer.....	68
4.3 Wahrscheinlichkeitsdichte am Ausgang des NF - Filters.....	99

5. Numerische Berechnung.....	102
5.1 Vorüberlegungen.....	102
5.2 Autokovarianzmatrix.....	103
5.3 Die komplexe Matrix A.....	104
5.4 Die Berechnung der Kehrmatrix und der Determinante...	104
5.5 Der Zentralitätsparameter.....	105
5.6 Abtastung und Fensterung.....	105
5.7 Duhamel-Hollmann-Split-Radix FFT.....	113
6. Simulation	118
6.1 Schätzung der Wahrscheinlichkeitsdichte.....	120
6.2 Normalverteilte Zufallszahlen.....	121
6.3 Erzeugung gleichverteilter Zufallszahlen.....	122
6.4 Lineare Kongruenzgeneratoren.....	126
6.5 Mehrfach rekursive Generatoren.....	131
6.6 Nichtlineare Kongruenzgeneratoren.....	132
7. Ergebnisse.....	138
8. Zusammenfassung.....	147
Literaturverzeichnis.....	150

Verzeichnis der benutzten Symbole

$\vec{H}_D$	Diagonalmatrix mit den Elementen des Vektors $\vec{h}$
$f(u)$	eindimensionale charakteristische Funktion der Zufallsvariablen $Y(i)$
$f_i^k(u)$	Näherung k-ter Ordnung für die eindimensionale charakteristische Funktion von $Y(i)$ zum Zeitpunkt i
$f_{x_i}^{\vec{x}}(u) = f(u_1, u_2, \dots, u_k) \dots$	k-dimensionale charakteristische Funktion des Zufallsvektors $\vec{X}_i$
$F(x)$	eindimensionale Verteilungsfunktion
$F_{x_i}^{\vec{x}}(x) = F(x_1, x_2, \dots, x_k) \dots$	k-dimensionale Verteilungsfunktion des Zufallsvektors $\vec{X}_i$
$g(n), h(n)$	Impulsantworten zeitdiskreter zeitinvarianter linearer Filter
$\vec{h} = (h(k-1), h(k-2), \dots, h(0))$	Vektor mit Elementen, die den ersten k Werten der Impulsantwort ent- sprechen
i	Beobachtungszeitpunkt
$\vec{I}$	Einheitsmatrix
j	Imaginäre Einheit
k	Anzahl der Dimensionen, Ordnung der Schätzung

Λ	Kovarianzmatrix
$\lambda_{m,n}$	Elemente der Kovarianzmatrix
$p(y)$	eindimensionale Wahrscheinlichkeits- dichte der Zufallsvariablen $Y(i)$
$p_i^k(y)$	Näherung k-ter Ordnung für die ein- dimensionale Wahrscheinlichkeits- dichte der Zufallsvariablen $Y(i)$ zum Zeitpunkt i
$p_{x_i}^{\rightarrow}(x) = p(x_1, x_2, \dots, x_k) \dots$	k-dimensionale Wahrscheinlichkeits- dichte des Zufallsvektors $\vec{X}_i$
$R(i)$	Autokorrelationsfunktion
$r(i) = R(i) / R(0)$	Autokorrelationskoeffizient
$s(i)$	Deterministisches Eingangssignal
$\vec{s}_i = (s(i-k+1), \dots, s(i-1), s(i))$	Signalvektor
$\sigma^2 = R(0)$	Varianz
$\{X_i, i \in \mathbb{Z}\}$	Stochastischer Prozeß am Eingang des Filters
$\vec{X}_i$	Zufallsvektor des Eingangssignals zum Zeitpunkt i
$X(i)$	Zufallsvariable am Eingang des Filters
$x(i)$	Realisation der Zufallsvariablen $X(i)$
$\{Y_i, i \in \mathbb{Z}\}$	Stochastischer Prozeß am Ausgang des Filters

1. Einführung

Stochastische Methoden und Modelle spielen in der Nachrichtentechnik eine wichtige Rolle. Dies liegt zum einen daran, daß unvermeidbare Störungen, die in Form von Rauschen auftreten, als stochastische Prozesse angesehen werden können; andererseits wird auch oft die Nachricht selbst als ein zufälliges Ereignis aufgefaßt. Nach dem Entropiesatz der Nachrichtentechnik besteht ein Zusammenhang zwischen dem Informationsgehalt eines Ereignisses und der Wahrscheinlichkeit seines Auftretens.

Bevor stochastische Methoden auf die Signale angewendet werden können, muß jedoch zunächst ein Modell gefunden werden, das die statistischen Eigenschaften des tatsächlichen Signals möglichst genau beschreibt. Aus diesem Modell kann dann ein Verfahren zur Kodierung, Dekodierung, Vorhersage des Signals oder Verbesserung des Signal-Rauschabstandes entwickelt werden.

Die Konstruktion eines passenden Modells ist dabei oft aufwendiger als der anschließende Entwurf eines Signalverarbeitungsverfahrens.

Sind die Eigenschaften von Signal- und Störgrößen erst einmal bekannt, so ist es oft wichtig, bestimmen zu können, wie sich deren stochastische Kenngrößen unter dem Einfluß von Signalverarbeitungsverfahren, wie lineare und nichtlineare Filterung, verändern.

Für viele statistische Kenngrößen ist dieses Problem bereits gelöst worden. So erhält man die Autokorrelationsfunktion am Ausgang eines linearen

Filters mit Hilfe der Wiener-Lee Formel, wenn die Autokorrelationsfunktion am Eingang des Filters bekannt ist.

Die Berechnung der Wahrscheinlichkeitsdichteverteilung nach linearer Filterung ist, vor allem für nicht gaußverteilte Eingangsgrößen, ein noch weitgehend ungelöstes Problem.

Ein typisches Anwendungsbeispiel, bei dem die Berechnung der Dichte nach linearer Filterung benötigt wird, ist die Abschätzung der Fehlerwahrscheinlichkeit in Detektorsystemen. Die Kenntnis der Varianz oder Autokorrelationsfunktion des Störprozesses reicht hierfür nicht aus, da verschiedene Dichteverteilungen bei gleicher Varianz zu sehr unterschiedlichen Fehlerwahrscheinlichkeiten führen.

Für den Fall analoger Detektoren und Filter ist dieses Problem bereits intensiv untersucht worden. Man erhielt allerdings Lösungen nur für wenige Spezialfälle, bei denen sowohl die Art der Verteilung als auch die verwendeten Filter bestimmten Bedingungen genügen mußten.

Hauptziel der vorliegenden Arbeit ist es, eine möglichst allgemeingültige Methode zur Berechnung der Wahrscheinlichkeitsdichteverteilung nach digitaler rekursiver und nichtrekursiver Filterung zu entwickeln.

Als Realisation eines nichtlinearen Filters sollen außerdem gedächtnisfreie Systeme mit quadratischer Kennlinie untersucht werden, da diese zum Bau von Detektorschaltungen benötigt werden.

2. Vorarbeiten

2.1. Vorarbeiten anderer Autoren

Klassische Methoden zur Berechnung der Wahrscheinlichkeitsdichte nach linearer Filterung beziehen sich auf analoge (kontinuierliche) Filter und Signale. Im Gegensatz dazu sollen in der vorliegenden Arbeit ausschließlich digitale (diskrete) Filter untersucht werden.

Zwei der klassischen Methoden werden dennoch vorgestellt, um zu zeigen, welche Schwierigkeiten bei der Berechnung der Wahrscheinlichkeitsdichte nach kontinuierlicher Filterung auftreten.

Eines dieser klassischen Probleme ist die Berechnung der Wahrscheinlichkeitsdichte für ein System, das aus einem Quadrierer und einem linearen Filter besteht (Bild 2.1). Diese Frage wurde zuerst von M. Kac und A. J. F. Siegert [34] untersucht.

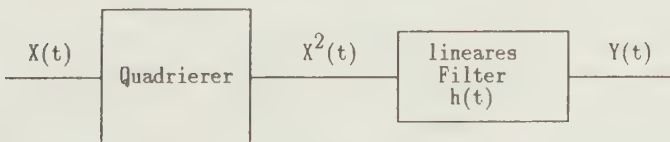


Bild 2.1 Quadrierer und Filter

Der Eingangsprozeß $\{X(t), t \in \mathbb{R}\}$ soll mittelwertfrei, normalverteilt und stationär sein und ist daher durch seine Autokorrelationsfunktion $R(\tau)$



vollständig beschrieben.

Den Ausgangsprozeß $\{Y(t), t \in \mathbb{R}\}$ erhält man durch Faltung des quadrierten Eingangsprozesses $\{X^2(t), t \in \mathbb{R}\}$ mit der Impulsantwort $h(t)$ des linearen Systems

$$Y(t) = \int_0^t h(t-\tau) \cdot X^2(\tau) d\tau \quad (2.1)$$

wobei $h(t)$ die Impulsantwort eines kausalen linearen Filters ist.

Kac und Siegert haben die Dichte von $Y(t)$ nun mit Hilfe einer Reihenentwicklung der charakteristischen Funktion $f_Y(u)$ entwickelt.

Aus ihr enthält man dann durch Fouriertransformation die Wahrscheinlichkeitsdichte. Umgekehrt ist die charakteristische Funktion die inverse Fouriertransformierte der Wahrscheinlichkeitsdichte oder der Erwartungswert der Funktion $\exp(juY(t))$ der Zufallsvariablen $Y(t)$.

$$\begin{aligned} f_Y(u) &= E [\exp \{juY(t)\}] \\ &= E [\exp \{ju \int_0^t h(t-\tau) \cdot X^2(\tau) d\tau\}] \end{aligned} \quad (2.2)$$

Kac und Siegert haben gezeigt, daß der Prozeß $\{Y(t), t \in \mathbb{R}\}$ als gewichtete Summe von statistisch unabhängigen normalverteilten Zufallsvariablen

G_i , $i \in \mathbb{N}$ dargestellt werden kann, mit

$$Y(t) = \sum_{i=1}^{\infty} \lambda_i G_i^2 \quad (2.3)$$

wenn $X(t)$ ein stationärer normalverteilter Prozeß mit endlicher Varianz ist. Die Konstanten λ_i in Gleichung (2.3) sind die reellen und diskreten Eigenwerte der homogenen Integralgleichung

$$\lambda_i \phi_i(\tau) = \int_0^{\infty} h(u) \cdot R(\tau-u) \cdot \phi_i(u) du \quad (2.4)$$

wobei $\phi_i(\tau)$ die zueinander orthogonalen Eigenfunktionen der Integralgleichung (2.4) sind und $R(\tau)$ die Autokorrelationsfunktion von $X(t)$ ist. Die statistisch unabhängigen Zufallsvariablen G_i sind dann

$$\lambda_i^{1/2} G_i = \int_0^t h(t-\tau) \cdot X(\tau) \cdot \phi_i(\tau) d\tau.$$

Mit Gleichung (2.3) vereinfacht sich die charakteristische Funktion des Ausgangsprozesses zu

$$f_y(u) = E [\exp \{ju \sum_{i=1}^{\infty} \lambda_i G_i^2\}]. \quad (2.5)$$

Wegen der statistischen Unabhängigkeit der normalverteilten Zufallsvariablen G_i kann man die charakteristische Funktion auch in Produktform schreiben:

$$f_y(u) = \prod_{i=1}^{\infty} E [\exp \{ju \lambda_i G_i^2\}]. \quad (2.6)$$

Die einzelnen Faktoren entsprechen dabei der charakteristischen Funktion von χ^2 -Verteilungen mit

$$\begin{aligned}
 f_y(u) &= E [\exp \{ju\lambda_i G_i^2\}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp \{ju\lambda_i x^2\} \cdot \exp \{-\frac{1}{2}x^2\} dx \\
 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp \{-\frac{1}{2}x^2(1 - 2ju\lambda_i)\} dx \\
 &= \frac{1}{\sqrt{1 - 2ju\lambda_i}}
 \end{aligned} \tag{2.7}$$

Die Wahrscheinlichkeitsdichte nach der Filterung ist dann die Fourier-Transformierte von (2.6)

$$\begin{aligned}
 p(y) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-juy) \cdot f_y(u) du \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-juy) \cdot \prod_{i=1}^{\infty} \frac{1}{\sqrt{1 - 2ju\lambda_i}} du
 \end{aligned} \tag{2.8}$$

In der Praxis stellt aber offenbar sowohl die Lösung der Integralgleichung (2.4) als auch die des Fourierintegrals ein kaum überwindbares Hindernis dar.

R. C. Emerson [25] hat deshalb vorgeschlagen, die Kumulanten an Stelle der charakteristischen Funktion zu bestimmen. Die Kumulanten K_s sind definiert durch

$$\ln(f(u)) = \sum_{s=1}^{\infty} \frac{K_s}{s!} (ju)^s \quad (2.9)$$

Die Wahrscheinlichkeitsdichte erhält man dann als Reihenentwicklung in Form von Gram-Charlier- und Laguerre- Polynomen direkt aus den Kumulanten, d.h. ohne Fourierinversion der charakteristischen Funktion. Die Kumulanten erhielt Emerson, indem er zunächst den Systemoperator der folgenden Anordnung bestimmte

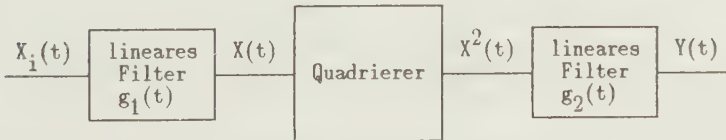


Bild 2.2 Quadrierer mit zwei linearen Filtern

Die Zufallsvariable nach dem ersten Filter ist dann

$$X(t) = \int_{-\infty}^{\infty} g_1(t-\tau) \cdot X_1(\tau) d\tau \quad (2.10)$$

und am Ausgang des Quadrierers haben wir

$$X^2(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(t-\tau) \cdot g_1(t-s) \cdot X_1(\tau) \cdot X_1(s) \, d\tau \, ds. \quad (2.11)$$

$Y(t)$ am Ausgang des Filters $g_2(t)$ ist dann

$$Y(t) = \int_{-\infty}^{\infty} g_2(t-\tau) \cdot X^2(\tau) \, d\tau. \quad (2.12)$$

Durch Einsetzen von Gleichung (2.11) in Gleichung (2.12) und Substitution der Variablen erhielt Emerson die Gleichung

$$Y(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X_1(t-u) \cdot h(u,v) \cdot X_1(t-v) \, du \, dv. \quad (2.13)$$

mit

$$h(u,v) = \int_{-\infty}^{\infty} g_1(u-z) g_2(z) g_1(v-z) \, dz. \quad (2.14)$$

Die rechte Seite von Gleichung (2.13) ist dabei der Systemoperator und $h(u,v)$ der Systemkern. Dieser Kern kann ebenfalls als Reihe dargestellt werden.

$$h(u,v) = \sum_{i=0}^{\infty} \lambda_i \cdot a_i(u) \cdot a_i(v) \quad (2.15)$$

wobei $a_i(u)$ und λ_i die Eigenfunktion bzw. Eigenwerte der Integralgleichung

$$\lambda_i a_i(u) = \int_{-\infty}^{\infty} h(u,v) \cdot a(v) dv \quad (2.16)$$

sind. Damit haben wir wieder eine homogene Integralgleichung eines ähnlichen Typs wie bei Kac und Siegert.

Die Lösungen der Gleichung (2.16) fand Emmerson allerdings nur unter der Bedingung, daß die Filter g_1 und g_2 eine gaußförmige Übertragungsfunktion mit

$$G_1(f) = \exp \left[-\frac{(f+f_0)^2}{2 \cdot b_1^2} \right] + \exp \left[-\frac{(f-f_0)^2}{2 \cdot b_1^2} \right] \quad (2.17)$$

und

$$G_2(f) = \exp \left[-\frac{f^2}{b_2^2} \right] \quad (2.18)$$

besitzen, wobei b_1 und b_2 die Bandbreiten der entsprechenden Filter sind. Diese Annahme hat allerdings zur Folge, daß die Filter g_1 und g_2 nicht kausal und damit auch nicht physikalisch realisierbar sind.

M.A. Meyer und D. Middleton [49] berechneten auf der Grundlage des Verfahrens von Kac und Siegert die Dichte am Ausgang eines Systems wie in Bild (2.2.) für verschiedene Grenzfälle der Bandbreite des Filters g_2 :

1. Ist das Filter g_2 des Systems breitbandig gegenüber dem Filter g_1 , so erhält man näherungsweise eine Dichte, die der vor dem Filter g_1 entspricht. Wegen des normalverteilten Eingangssignals und des Quadrierers ist dies eine χ^2 -Verteilung mit einem Freiheitsgrad.

2. Ist die Bandbreite von $g_2(t)$ sehr klein gegenüber $g_1(t)$, so erhält man näherungsweise eine Normalverteilung am Ausgang des Systems.

Die Berechnung der Wahrscheinlichkeitsdichte von stochastischen Prozessen nach digitaler Filterung scheint – im Vergleich zum analogen Fall – noch wenig erforscht zu sein.

H. W. Schüller [56] schrieb zu diesem Thema:

"Die Frage liegt nahe, wie sich die Verteilungsdichte eines Zufallsprozesses unter dem Einfluß eines linearen Systems verändert. Hier ist eine verhältnismäßig einfache Aussage nur möglich, wenn die einzelnen Werte der Eingangsfolge voneinander statistisch unabhängig sind. Die Ausgangsfolge des Systems

$$Y(k) = \sum_{\kappa} h_0(\kappa) \cdot V(k-\kappa)$$

ist dann die Summe der voneinander unabhängigen Elemente der Eingangsfolge nach ihrer Gewichtung mit den $h_0(\kappa)$."

Wegen der statistischen Unabhängigkeit erhält man die Wahrscheinlichkeitsdichte $p(y)$ von $Y(t)$, indem man die mit $h_0(\kappa)$ gewichteten Wahrscheinlichkeitsdichten $p_x(x)$ der Zufallsvariablen am Eingang miteinander faltet.

$$p(y) = \prod_i^* p_i(x) = p_1(x) * p_2(x) * p_3(x) \dots \quad (2.19)$$

mit

$$p_i(x) = p_x(x/h_0(i)) / |h_0(i)| \quad (2.20)$$

Bei statistischer Unabhängigkeit erhält man auch für die charakteristische Funktion eine sehr einfache Lösung. Die charakteristische Funktion der Summe von statistisch unabhängigen Zufallsvariablen ist nämlich das Produkt der charakteristischen Funktionen der einzelnen Zufallsvariablen.

Zur Berechnung der charakteristischen Funktion nach der Filterung müssen diese dann wieder mit $h_0(\kappa)$ gewichtet werden. Bei H. W. Schübler [56] finden wir die Formel

$$f_y(u) = \prod_i f_x(u \cdot h(i)) \quad (2.21)$$

Allerdings dürfte das Faltungsprodukt für die Wahrscheinlichkeitsdichte bzw. die Fouriertransformierte der charakteristischen Funktion in den seltensten Fällen geschlossen bestimmbar sein.

Schüßler beschränkte sich daher auf die Berechnung der Wahrscheinlichkeitsdichte einer Linearkombination von zwei normalverteilten Zufallsvariablen.

Die Gleichungen (2.19) und (2.21) sind, wie bereits erwähnt, nur für statistisch unabhängige Eingangsfolgen gültig und können nicht ohne weiteres auf andere Fälle erweitert werden. Das Ziel der vorliegenden Arbeit ist im Gegensatz dazu die Entwicklung eines möglichst allgemeingültigen Verfahrens zur Bestimmung der Dichte nach linearer Filterung, wobei beliebige Eingangsverteilungen und statistische Abhängigkeiten zugelassen sein sollen.

2.2 Eigene Vorarbeiten

Die vorliegende Arbeit wurde aus einer Fragestellung entwickelt, die im Arbeitsvorhaben Sprachverarbeitung im Sommersemester 1984 auftauchte. Damals wurde vorgeschlagen, zur Verbesserung der Sprachverständlichkeit nicht nur das Leistungsdichtespektrum, sondern auch die Wahrscheinlichkeitsdichte bei der Synthese von Sprachsignalen zu berücksichtigen, wobei noch ungeklärt ist, ob diese überhaupt hörbar ist. Dazu sollte die Sprachquelle (Quelle-Filtermodell) die noch zu messende Verteilung des natürlichen Lautes erhalten, ein nachgeschaltetes Filter sollte dann für das passende Leistungsdichtespektrum sorgen. Prof. Dr. Wasiljeff machte uns damals darauf aufmerksam, daß die Wahrscheinlichkeitsdichte durch lineare Filterung verändert wird. Die Frage, wie diese nach linearer Filterung zu berechnen ist, konnte im Rahmen dieses Arbeitsvorhabens jedoch nicht geklärt werden.

Zur Lösung des Problems der linearen Filterung stochastischer Prozesse wurde zunächst eine Anordnung bestehend aus einer Rauschquelle und einem diskreten digitalen Filter simuliert.

Der Einfachheit halber wurde dabei eine Quelle mit statistisch unabhängigen identisch verteilten Abtastwerten und als Filter ein digitales nichtrekursives Filter verwendet.

Im Rahmen zweier Diplomarbeiten wurden auch theoretische Lösungen für die Wahrscheinlichkeitsdichte der simulierten Anordnung gefunden.

indem das Problem auf die Berechnung der Wahrscheinlichkeitsdichte von Summen endlich vieler nichtidentisch verteilter unabhängiger Zufallsvariablen zurückgeführt wurde.

Wie sich später herausstellte, entspricht dies im wesentlichen der von Schüßler vorgeschlagenen Methode.

Einen neuen Beitrag lieferten beide Diplomarbeiten aber dennoch, da die diskrete Faltung zur Berechnung des Faltungsprodukts (Mantei, Schreiber [43]) und die diskrete Fouriertransformation (Simmer, Spöring [58]) zur numerischen Bestimmung der Wahrscheinlichkeitsdichte eingesetzt wurden.

Hauptnachteil der in den Diplomarbeiten vorgestellten Methode ist, daß sie nur den Fall statistisch unabhängiger Abtastwerte am Eingang digitaler nichtrekursiver Filter berücksichtigen.

Um die Methode auch auf statistisch abhängige Zufallsvariablen zu erweitern, wurde zunächst nach einer Quelle gesucht, die eine solche Zeitreihe bildet, und dann versucht, die Verteilung nach linearer Filterung zu bestimmen. Da das Verhalten auf einem Digitalrechner simuliert werden sollte, bot es sich an, einen bereits vorhandenen Schieberegister-

Zufallszahlengenerator näher zu untersuchen.

Der benutzte Generator (Bild 2.3) besteht aus einem Register mit 16 Speicherzellen mit vier Anzapfungen, an die zwei EXOR - Gatter angeschlossen sind. Bei jedem Taktimpuls geben nun die Speicherzellen ihren Inhalt an den linken Nachbarn weiter, falls dieser vorhanden ist. Zusätzlich wird der Inhalt der Speicher 15 und 13 sowie 12 und 10 auf ein EXOR - Gatter geführt. Die Ausgänge dieser Gatter sind mit einem weiteren EXOR -

Gatter verbunden, dessen Ausgang dann auf die erste Speicherzelle rückgekoppelt wird.

Solche Generatoren werden häufig benutzt, um Folgen von "zufälligen" binären Werten, d.h. -1 oder 1 , zu produzieren.

Man liest dazu bei jedem Taktimpuls jeweils nur den Wert von einer bestimmten Speicherzelle, z.B. Zelle 0, aus.

Die "statistischen" Eigenschaften solcher Sequenzen hängen dabei von der Lage der Anzapfungen ab. Sind diese gut gewählt, so kann die entstehende Folge näherungsweise als statistisch unabhängig ansehen werden. Man spricht dann von einem Pseudo-Rauschgenerator.

Besteht ein solches rückgekoppeltes Schieberegister aus 16 Speichern, so kann man den Inhalt auch als 16-stellige Dualzahl auffassen (Für längere Register gilt sinngemäß das gleiche). Man erhält dann nach geeigneter Normierung Zufallszahlen, die auf dem Intervall $(-1,1)$ gleichverteilt sind.

Man kann rückgekoppelte Schieberegister also auch als Pseudozufallszahlengeneratoren verwenden.

Allerdings sind diese Zahlen voneinander abhängig. Besteht ein solches Register aus 16 Speicherzellen, so erzeugt es Zahlen, von denen jeweils 16 "statistisch" abhängig sein müssen, d.h. zur Erzeugung "statistisch" unabhängiger Zahlen darf man nur jede 16-te Zahl auswerten, wodurch die Methode recht langsam wird. Für unsere Zwecke ist es jedoch interessant, einmal zu untersuchen, wie sich eine Zahlenreihe verhält, wenn man nach jeder Schiebeoperation eine Zahl entnimmt. Wir erhalten gleichverteilte Zufallszahlen, die "statistisch" abhängig sind und daher ein ungleichförmiges Leistungsdichtespektrum besitzen (Bild 2.4).

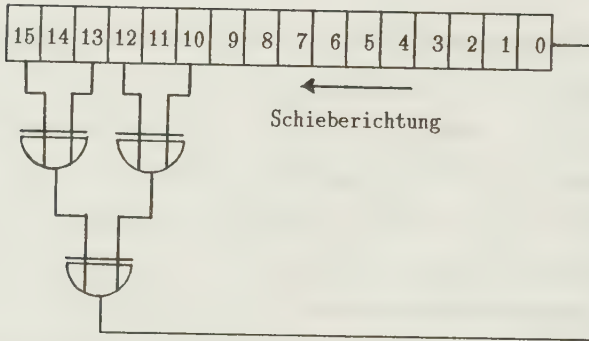


Bild 2.3 Ein rückgekoppeltes Schieberegister

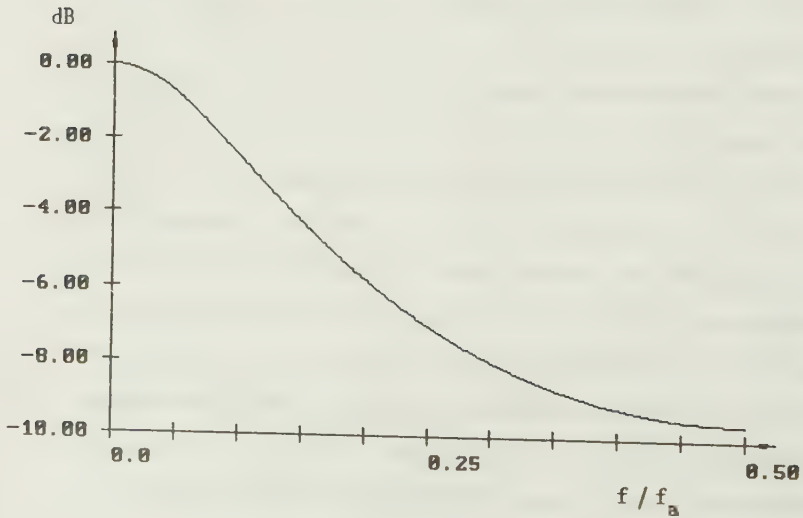


Bild 2.4 Leistungsdichtespektrum des Generators

Aus eigenen Versuchen war aber bekannt, daß man ein mit graphischer Genauigkeit übereinstimmendes Leistungsdichtespektrum erhält, wenn man eine statistisch unabhängige Quelle mit einem rekursiven Filter mit der folgenden Impulsantwort filtert:

$$q(i) = 2^{(i-N)} \quad i = 0, 2, \dots, N-1 \quad (2.22)$$

N ist dabei die Anzahl der Speicherzellen. Es drängte sich deshalb die Vermutung auf, daß das Schieberegister möglicherweise als Kombination aus einer statistisch unabhängigen Quelle – mit zunächst noch unbekannter Verteilung – und einem linearen Filter aufgefaßt werden könnte.

Bei der Konstruktion eines solchen Modells treten aber vor allem zwei Schwierigkeiten auf: Zum einen ist die Verteilung am Ausgang des Generators (Gleichverteilung) bekannt, nicht aber die Verteilung der Quelle, zum anderen können die Filterkoeffizienten nicht aus dem Leistungsdichtespektrum bestimmt werden, da die Phase unbekannt ist. Außerdem sind Filter mit Nullstellen nicht invertierbar.

Zur Bestimmung der Wahrscheinlichkeitsdichte von gefilterten Schieberegistersignalen ist das Quelle-Filter Modell aber gut geeignet.

Durch Invertierung der Übertragungsfunktion des vermuteten Quellenfilters erhielten wir eine diskrete Wahrscheinlichkeitsdichte, die nur bei -1 und 1 einen von Null verschiedenen Wert hat. (Bild 2.5).

Das beschriebene Quelle – Filtermodell kann aber auch direkt aus dem Erzeugungsmechanismus des Schieberegister – Pseudo – Zufallszahlen-generators abgeleitet werden: Quelle ist dann ein Pseudo – Rausch-

generator, der nur binäre Ziffern -1 und $+1$ produziert. Die Bildung von Zahlen aus jeweil 16 dieser Ziffern kann dann als Filterung aufgefaßt werden, da diese bei jedem Taktimpuls nur um eine Stelle verschoben und mit den Koeffizienten $q(i)$ aus Gleichung (2.22) gewichtet werden. Daraus ergibt sich durch inverse Fouriertransformation die charakteristische Funktion der Quelle zu

$$f(u) = \cos(u) \quad (2.23)$$

Filtert man die Zahlenfolge des Schieberegister-Pseudo-Zufallszahlengenerators erneut, so kann man das (gedachte) Filter $q(i)$ der Quelle und das äußere Filter $h(i)$ zu einem einzigen Filter $g(i)$ zusammenfassen.

Die Impulsantwort des Gesamtfilters erhält man dann durch Faltung der einzelnen Impulsantworten. (Bild 2.6)

Liefert die Quelle (ohne das gedachte Filter) statistisch unabhängige Zahlenfolgen, dann kann man das bereits beschriebene Verfahren zur Bestimmung der Wahrscheinlichkeitsdichte nach der Filterung anwenden.

Mit der Produktformel für die charakteristische Funktion haben wir

$$f_y(u) = \prod_i^N f_x(u \cdot g(i)) \quad (2.24)$$

mit

$$g(i) = q(i) * h(i) \quad (2.25)$$

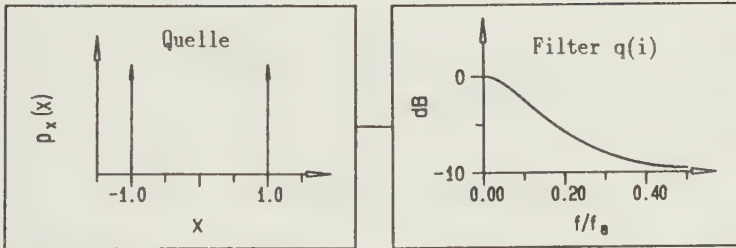


Bild 2.5 Quelle-Filter Modell für das Schieberegister

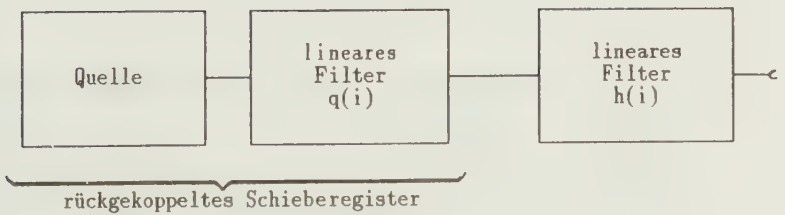


Bild 2.6 Filterung des Schieberegistersignals

Für ein Filter mit der Impulsantwort $h(1) = h(2) = 0.5$ erhalten wir für Gleichung (2.25)

$$g(i) = \frac{3}{8} \cdot 2^{(i+1-N)} \quad i = 0, 1, \dots, N-1$$

und

$$g(N) = \frac{1}{4} \quad (2.26)$$

Dies führt auf die in Bild 2.7 gezeigte Wahrscheinlichkeitsdichte.

Das Ergebnis konnte durch eine Simulation mit 10000 Abtastwerten (Bild 2.8) bestätigt werden.

Diese Methode wurde auf der EUSIPCO-86 vorgestellt [44]. Bei der anschließenden Diskussion des Aufsatzes stellte sich aber heraus, daß dieser Weg nicht beliebig verallgemeinert werden kann. Nicht jedes Rauschen läßt sich in eine statistisch unabhängige Quelle und ein Filter zerlegen, im Gegenteil, dies stellt einen seltenen Spezialfall dar. Für die wichtige Klasse der nichtstationären Prozesse ist eine Zerlegung in keinem Fall möglich.

Diese Kritik lieferte den Anstoß für die vorliegende Arbeit.

Hier soll vor allem die Frage geklärt werden, aus welchen statistischen Kenngrößen des Eingangssignals die Ausgangsverteilung eindeutig bestimmt werden kann. Für das Problem, das Kac und Siegert untersuchten, war dies die Autokorrelationsfunktion $R(\tau)$ eines stationären, normalverteilten Eingangsprozesses. Bei der zuletzt beschriebenen Methode wird außer der Wahrscheinlichkeitsdichte $p(x)$ noch das Leistungsdichtespektrum benötigt. Beide Ergebnisse können aber nicht verallgemeinert werden.

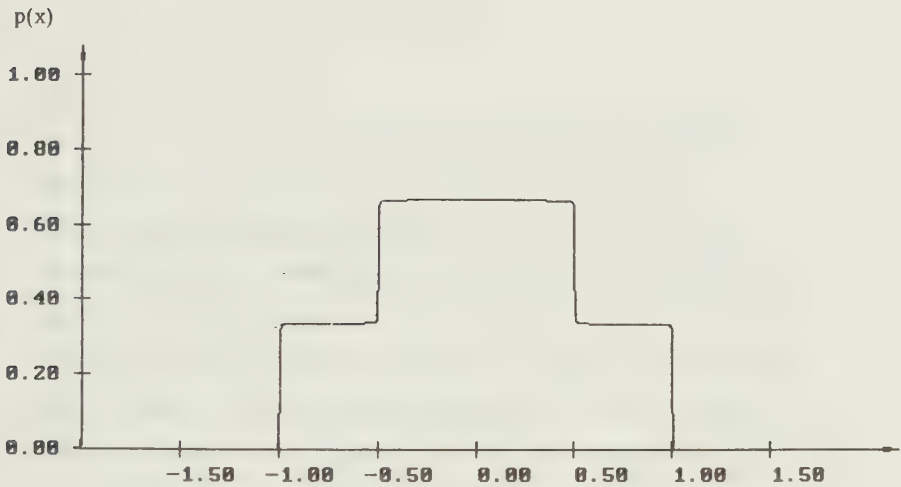


Bild 2.7 Theoretische Wahrscheinlichkeitsdichte nach der Filterung

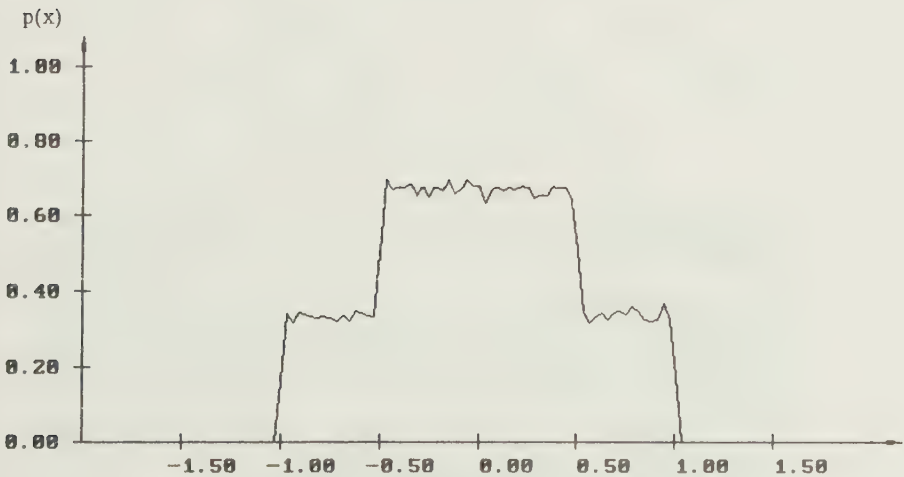


Bild 2.8 Simulationsergebnis für die Wahrscheinlichkeitsdichte

3. Filterung stochastischer Prozesse

Untersucht werden soll nun eine Anordnung bestehend aus einer Rauschquelle und einem linearen digitalen Filter. Die Quelle erzeugt einen diskreten stochastischen Prozeß, der stationär oder nichtstationär sein kann und aus Zufallsvariablen besteht, die auf irgend eine Weise statistisch abhängig sind und eine beliebige Wahrscheinlichkeitsdichteverteilung besitzen. Von besonderem Interesse sind dabei Rauschquellen, deren Zufallsvariablen nicht gaußverteilt sind. Gesucht wird ein Verfahren zur Berechnung der Wahrscheinlichkeitsdichte der Zufallsvariablen $Y(i)$ am Ausgang des Filters.

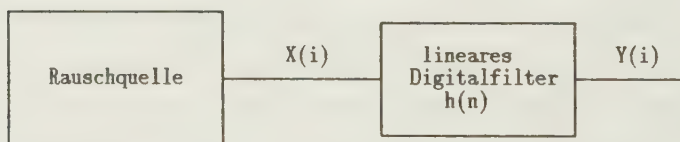


Bild 3.1 Filterung von Rauschen

Als Filtereigenschaften setzen wir Linearität, Zeitinvarianz, Stabilität und Kausalität voraus.

Nichtrekursive und rekursive Filter mit beliebiger Struktur sind zugelassen; Grenzyklen und Eigenrauschen der Filters werden jedoch nicht berücksichtigt.

Aus mathematischer Sicht ist die Wirkung eines solchen Filters die Anwendung eines Faltungsoperators auf das Eingangssignal.

Man erhält das Ausgangssignal, indem man den Eingangsprozeß mit der Eingangsimpulsantwort des Systems faltet.

Zur Berechnung des Ausgangssignals bei analoger (kontinuierlicher) Filterung muß man dazu das Faltungsintegral

$$Y(t) = \int_{-\infty}^{\infty} h(t-\tau) \cdot X(\tau) d\tau \quad (3.1)$$

lösen. Die Berechnung der Wahrscheinlichkeitsdichte am Ausgang erfordert, wie wir in Kapitel 2.1 gesehen haben, die Lösung der homogenen Integralgleichung (2.4)

$$\lambda_1 \phi_1(\tau) = \int_0^{\infty} h(u) \cdot R(\tau-u) \cdot \phi_1(u) du. \quad (3.2)$$

Eine Lösung ist bisher aber nur für wenige Spezialfälle bekannt.

Im Fall digitaler diskreter Filter vereinfacht sich das Faltungsintegral zu einer Faltungssumme.

Diese Betrachtungsweise ermöglicht es, sich von der physikalischen Filterstruktur vollkommen zu lösen, was für die weiteren Überlegungen ein großer Vorteil ist.

Man erhält dann die Zufallsvariablen $Y(i)$ des Ausgangsprozesses als gewichtetes Mittel der Variablen $X(i)$ des Eingangsprozesses. Die Berechnung der Wahrscheinlichkeitsdichte solcher Prozesse ist, wie wir noch sehen werden, in in sehr viel allgemeingültigerer Form möglich als für den Fall analoger Filter.

Schwierigkeiten entstehen allerdings dadurch, daß die Anzahl der Summanden im allgemeinen Fall unendlich groß ist; die Gewichtung der Zufallsvariablen nimmt aber bei stabilen Filtern mit dem Ausschwingen der Impulsantwort schnell ab.

3.1 Rauschen und stochastische Prozesse

Als Rauschen bezeichnet man in der Nachrichtentechnik Störungen, die als zufällige Vorgänge aufgefaßt werden können (engl. random noise). Die Spannung zu einem festen Zeitpunkt t ist dann die Realisierung einer Zufallsvariablen. Betrachtet man die Entwicklung des Rauschens über die Zeit t , so ist die mathematische Beschreibung des Vorgangs der sogenannte stochastische Prozeß. Dieser ist nach Fisz ([26]S. 320) wie folgt definiert:

Definiton 5.1. Eine Menge von Zufallsvariablen $X(t)$, die von einem Parameter t abhängen, der in einer gewissen Zahlenmenge I variiert, nennen wir einen stochastischen Prozeß und bezeichnen ihn mit $\{X_t, t \in I\}$. Der

Parameter t ist dabei die Zeit oder eine andere Größe mit oder ohne physikalische Bedeutung.

Da wir uns mit der digitalen und damit zeitdiskreten Filterung von stochastischen Prozessen beschäftigen wollen, erhalten wir als Realisation des Prozesses nur zu bestimmten Zeitpunkten $t \in t_1, t_2, \dots$ einen von Null verschiedenen Wert. Man bezeichnet dies als stochastischen Prozeß mit diskreter Zeit. Die Differenz Δt zwischen zwei aufeinander folgenden Zeitpunkten t_i, t_{i+1} nehmen wir für die weiteren Überlegungen als konstant an; ungleichförmige oder zufällige Abstände [5] werden in der Nachrichtentechnik gelegentlich aber auch benutzt. In unserem Fall gilt

$$t_i = i \cdot \Delta t \quad i \in \mathbb{Z}. \quad (3.3)$$

zur Vereinfachung kann man auch normierte Zeiten benutzen, indem man durch Δt dividiert

$$t_{\text{norm}} = t_i / \Delta t = i \quad \text{mit } i \in \mathbb{Z} \quad (3.4)$$

Wir haben nun einen stochastischen Prozeß mit der Indexmenge $\mathbb{Z}$:

$$\{X_i, i \in \mathbb{Z}\}$$

Im folgenden werden wir nur noch Teilmengen von stochastischen Prozessen betrachten. Enthält die Parametermenge I dieses Ausschnitts mehr als ein Element, so ist diese Teilmenge nach Definition 3.1 ebenfalls ein stochastischer Prozeß.

3.2. Lineare Filter

Diskrete lineare und zeitinvariante Filter sind, wie bereits erwähnt, durch ihre Impulsantwort $h(n)$ vollständig beschrieben. Ist i der Beobachtungszeitpunkt und $\{X_i, i \in \mathbb{Z}\}$ der Eingangsprozeß, dann erhalten wir die Zufallsvariablen des Ausgangsprozesses $\{Y_i, i \in \mathbb{Z}\}$ als Faltungssumme

$$Y(i) = \sum_{n=-\infty}^{\infty} h(n) \cdot X(i-n) \quad (3.5)$$

Gesucht ist nun die Wahrscheinlichkeitsdichte $p(y)$ der Zufallsvariablen $Y(i)$. Diese ist bei nichtstationären Prozessen vom Beobachtungszeitpunkt i abhängig, bei stationären Prozessen von i unabhängig.

Physikalisch realisierbare Filter sind stets kausal, d.h. die Impulsantwort $h(n)$ ist Null für $k < 0$. Die Faltungssumme ist dann

$$Y(i) = \sum_{n=0}^{\infty} h(n) \cdot X(i-n) \quad (3.6)$$

Außerdem haben wir Stabilität vorausgesetzt.

Von stabilen Filtern spricht man dann, wenn die Antwort des Systems auf ein Eingangssignal, das zu allen Zeitpunkten beschränkt ist, ebenfalls zu allen Zeitpunkten beschränkt ist.

Eine hinreichende und notwendige Bedingung für Stabilität ist die Beschränktheit der Beträge von $h(n)$, d.h.

$$\sum_{n=-\infty}^{\infty} |h(n)| = M < \infty. \quad (3.7)$$

Daß die Bedingung (3.7) hinreichend ist, kann man durch die folgende Abschätzung zeigen ([2], S.78):

$$|y(i)| \leq \sum_{n=-\infty}^{\infty} |h(n)| \cdot |x(i-n)| \quad (3.8)$$

unter der Voraussetzung der Beschränktheit der Eingangszahlenfolge $|x(i)| < K$ und der Bedingung (3.7) folgt daraus

$$|y(i)| \leq K \cdot \sum_{n=-\infty}^{\infty} |h(n)| = K \cdot M < \infty. \quad (3.9)$$

Damit nimmt die Folge $y(i)$ unabhängig vom Beobachtungszeitpunkt i nur endliche Werte an.

Daß die Bedingung (3.7) auch notwendig ist, ergibt sich aus dem folgenden Beispiel ([55], S.12):

Ist die Summe der Beträge der Impulsantwort nicht beschränkt, dann liefert die beschränkte Eingangsfolge

$$x(i) = \operatorname{sgn}(h(i))$$

$$\text{mit} \quad \operatorname{sgn}(h(i)) = \begin{cases} +1 & \text{für } h \geq 0 \\ -1 & \text{für } h < 0 \end{cases} \quad (3.10)$$

eine für $n \rightarrow \infty$ eine über alle Schranken wachsende Ausgangsfunktion.

Eine direkte Folge der Stabilitätsbedingung (3.7) ist, daß Filter mit dieser Eigenschaft für $n \rightarrow \infty$ ausschwingen.

Der Maximalwert des Betrags von $h(n)$ kann nämlich abgeschätzt werden mit

$$|h(i)| \leq \sum_{n=i}^{\infty} |h(n)| = \sum_{n=-\infty}^{\infty} |h(n)| - \sum_{n=-\infty}^{i-1} |h(n)|. \quad (3.11)$$

Aus (3.7) folgt dann

$$|h(i)| \leq M - \sum_{n=-\infty}^i |h(n)|. \quad (3.12)$$

Da die Summe $\sum_{n=-\infty}^i |h(n)|$ eine monoton steigende Funktion von i ist, die gegen M konvergiert, gilt damit

$$\lim_{n \rightarrow \infty} h(n) = 0 \quad (3.13)$$

Die Stabilitätseigenschaft ist deshalb wichtig, weil die Bestimmung der Wahrscheinlichkeitsdichte $p(y)$ der Zufallsvariablen $Y(i)$ erhebliche Probleme aufwirft, wenn die Impulsantwort des Filters $h(n)$ unendlich lang ist; und dies ist im allgemeinen der Fall.

Am Eingang des Filters liegt dann zu jedem Zeitpunkt i der vollständige stochastische Prozeß $\{X_i, i \in \mathbb{Z}\}$ an wie Gleichung (3.5) zeigt.

Bei kausalen Filtern formt der Faltungsoperator aus allen Zufallsvariablen, die auf der negativen Halbgeraden der Zeitachse (der Vergangenheit) liegen, zu jedem Zeitpunkt i jeweils eine neue Zufallsvariable $Y(i)$.

In beiden Fällen haben wir es mit unendlich vielen, und im allgemeinen statistisch abhängigen, Zufallsvariablen zu tun.

Da die Form der Abhängigkeit die Wahrscheinlichkeitsdichte am Ausgang beeinflußt, benötigen wir demnach eine Aussage über die statistischen

Abhängigkeiten zwischen allen Zufallsvariablen des stochastischen Prozesses am Eingang.

Eine wesentliche – und für die weiteren Überlegungen entscheidende – Vereinfachung erhalten wir, indem wir die Stabilitätseigenschaft und das damit verbundene asymptotische Verhalten $\lim_{n \rightarrow \infty} h(n) = 0$ der Impulsantwort ausnutzen.

Dazu definieren wir eine Näherung k -ter Ordnung für die Zufallsvariable $Y(i)$ zu

$$Y_i^k = \sum_{n=0}^{k-1} h(n) \cdot X(i-n) \quad (3.14)$$

indem wir die Impulsantwort bei $n=k$ abbrechen.

Wir bestimmen nun die Wahrscheinlichkeitsdichte der Zufallsvariablen Y_i^k und erhalten die Wahrscheinlichkeitsdichte von $Y(i)$ durch den Grenzübergang $k \rightarrow \infty$.

Für ein festes k hat diese Art der Darstellung den Vorteil, daß zu jedem Zeitpunkt i nur eine, auf k Zufallsvariable begrenzte Teilmenge des Eingangsprozesses $\{X_i, i \in \mathbb{Z}\}$ am Filter anliegt.

Für $k \leq 2$ ist diese Teilmenge ebenfalls ein stochastischer Prozeß

$$\{X_i, i \in (i-k+1, i-k+2, \dots, i-1, i)\}. \quad (3.15)$$

Da es sich um einen Prozeß mit diskreter Zeit handelt, ist aber auch eine andere Interpretation möglich. Der Prozeß am Eingang ist nichts anderes als eine Folge von k Zufallsvariablen

$$X(i-k+1), X(i-k+2), \dots, X(i-1), X(i)$$

Eine Folge von Zufallsvariablen kann als eine mehrdimensionale Zufallsvariable aufgefaßt werden. Die Realisation dieser Zufallsvariablen ist dann eine Folge von k reellen Zahlen. Diese reelle Zahlenfolge kann man geometrisch interpretieren; sie definiert einen Punkt im k -dimensionalen euklidischen Raum R_k . Die Zufallsvariable, die aus der Folge von Zufallsvariablen gebildet wurde, wird daher mehrdimensionale Zufallsvariable oder Zufallsvektor genannt.

Der k -dimensionale Zufallsvektor $\hat{X}_i$ am Eingang des Filters besteht also aus den Elementen

$$\hat{X}_i = (X(i-k+1), X(i-k+2), \dots, X(i-1), X(i)) \quad (3.16)$$

wobei i wieder der Beobachtungszeitpunkt ist. Zu einer festen Zeit i haben wir am Ausgang aber nur eine einfache (eindimensionale) Zufallsvariable. Aus geometrischer Sicht kann man die Wirkung eines Filters deshalb wie folgt beschreiben:

Ein Filter ist eine Funktion, die einen k -dimensionalen, im allgemeinen Fall sogar unendlich-dimensionalen Ereignisraum auf einen eindimensionalen abbildet. Die Anzahl der Dimensionen entspricht dabei der Länge der Impulsantwort des Filters.

Anders ausgedrückt bildet ein Filter zu jedem Zeitpunkt einen k - dimensionalen Zufallsvektor auf eine eindimensionale Zufallsvariable ab.

Diese Betrachtungsweise scheint neu zu sein; sie ermöglicht es, Sätze aus der Theorie der Zufallsvektoren auf die Filterung anzuwenden.

Diese Eigenschaft der Filterfunktion hat außerdem wichtige Konsequenzen für die Berechnung der Wahrscheinlichkeitsdichte am Ausgang des Filters. Die gewöhnliche eindimensionale Dichte $p(x)$ des Eingangsprozesses determiniert die eindimensionale Wahrscheinlichkeitsdichte am Ausgang noch nicht; entscheidend ist die gemeinsame k -dimensionale Wahrscheinlichkeitsdichte der Elemente des Zufallsvektors.

Ob allerdings die Zufallsvariable Y_1^k gegen $Y(i)$ konvergiert, hängt neben den Filtereigenschaften natürlich auch von den statistischen Eigenschaften des Eingangsprozesses ab. Beispielsweise divergiert die Schätzung, wenn wir einen nichtstationären Prozeß haben, dessen Varianz schneller wächst, als die Impulsantwort ausschwingt. Die Güte der Näherung muß daher im Einzelfall geprüft werden.

In der Nachrichtentechnik dürften solche pathologischen Fälle aber kaum vorkommen, denn wenn Y_1^k für große k eine schlechte Näherung ist, bedeutet

dies, daß sehr kleine Werte der Impulsantwort nach Gleichung (3.12) einen großen Beitrag zum Ausgangsprozeß liefern müssen. Das Verhalten des Ausgangsprozesses hängt dann sehr stark von der physikalischen Realisation des Filters ab (z.B. Filterstruktur und Wortlänge bei der digitalen Multiplikation und Addition). Solche Prozesse werden wir deshalb im folgenden nicht weiter untersuchen.

Bei einer bestimmten Klasse von digitalen Filtern sind solche Schätzungen ohnehin unnötig, weil die Länge der Impulsantwort beschränkt ist. Solche Filter sind stets stabil.

In die Klasse gehören alle nichtrekursiven Filter. Es gilt dann

$$Y(i) = \sum_{n=0}^{N-1} h(n) \cdot X(i - n) \quad (3.17)$$

Das Filter also hat ein "Gedächtnis" der Länge N .

3.3 Charakteristische Funktion und Verteilungsfunktion

Bevor wir die Wahrscheinlichkeitsdichte nach der Filterung bestimmen, definieren wir zuerst die dazu benötigten ein- und mehrdimensionalen stochastischen Kenngrößen.

Definition 3.3.1

Die Verteilungsfunktion $F(x)$ einer eindimensionalen Zufallsvariablen ist definiert durch die Wahrscheinlichkeit, daß die Zufallsvariable X unter einer Schwelle x bleibt, oder durch die Formel

$$F(x) = P(X < x). \quad (3.18)$$

wobei x eine reelle Zahl ist. ([26], S. 51)

Ein Zufallsvektor aus k Elementen ist durch seine k -dimensionale Verteilungsfunktion vollständig beschrieben.

Definition 3.3.2

Als k -dimensionale Verteilungsfunktion $F_x^*(\vec{x})$ des Zufallsvektor $\vec{X}_1$ mit den Elementen $(X_1, X_2, \dots, X_k)$ definieren wir die reelle Funktion

$$F_x^*(\vec{x}) = F(x_1, x_2, \dots, x_k) = P(X_1 < x_1, X_2 < x_2, \dots, X_k < x_k) \quad (3.19)$$

wobei $\vec{x}$ ein Vektor mit reellen Elementen ist. ([26], S. 62)

Zur Definition der Wahrscheinlichkeitsdichte führen wir den Begriff der stetigen Zufallsvariablen ein.

Definition 3.3.3

Eine Zufallsvariable X ist stetig, wenn eine Funktion $p(x) \geq 0$ existiert, die für jedes reelle x die Beziehung

$$F(x) = \int_{-\infty}^x p(x) \, dx \quad (3.20)$$

erfüllt, wobei $F(x)$ die Verteilungsfunktion von X ist. Die Funktion $p(x)$ nennen wir Wahrscheinlichkeitsdichte der Zufallsvariablen X . ([26], S. 54)

Wir erhalten die Wahrscheinlichkeitsdichte $p(x)$ daher durch Differenzieren der Verteilungsfunktion. Ist $F(x)$ im Punkt x stetig, so erhalten wir

$$p(x) = F'(x) \quad (3.21)$$

Die k -dimensionale Wahrscheinlichkeitsdichte $p_{\vec{x}}(\vec{x})$ ist festgelegt durch die k -dimensionale Verteilungsfunktion.

Definition 3.3.4

Eine k -dimensionale Zufallsvariable oder Zufallsvektor heißt stetig, wenn eine nichtnegative Funktion $p_{\vec{x}}(\vec{x}) = p_{\vec{x}}(x_1, x_2, \dots, x_k)$ existiert, die für jeden reellen Vektor $\vec{x}$ die Beziehung

$$F(\vec{x}) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_k} p(\vec{x}) dx_1 dx_2 \dots dx_k \quad (3.22)$$

erfüllt. $F(\vec{x})$ ist dabei Verteilungsfunktion des Zufallsvektors $\vec{X}$. ([26], S. 64.)

Ist die Wahrscheinlichkeitsdichte im Punkt $\vec{x}$ stetig, dann ist die Wahrscheinlichkeitsdichte ([41], S. 13)

$$p_{\vec{x}}(\vec{x}) = p_{\vec{x}}(x_1, x_2, \dots, x_k) = \frac{\partial^k F(x_1, x_2, \dots, x_k)}{\partial x_1 \partial x_2 \dots \partial x_k}. \quad (3.23)$$

Wie wir noch sehen werden, läßt sich die Wahrscheinlichkeitsdichte nach linearer Filterung besonders einfach mit Hilfe der charakteristischen Funktion bestimmen. Diese kann man als einen Erwartungswert einer bestimmten Funktion $g(x)$ der Zufallsvariablen X auffassen. Wir legen deshalb zuerst fest, was wir unter einem Erwartungswert verstehen.

Definition 3.3.5

Ist $p(x)$ die Wahrscheinlichkeitsdichte der Zufallsvariablen X , $g(x)$ eine integrierbare Funktion und die Ungleichung

$$\int_{-\infty}^{\infty} |g(x)| \cdot p(x) \, dx < \infty \quad (3.24)$$

erfüllt, dann nennen wir das Integral

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot p(x) \, dx \quad (3.25)$$

den Erwartungswert der Zufallsvariablen $g(X)$. ([26], S. 87)

Mit $g(X) = \exp(juX)$ geht dieser in die charakteristische Funktion über.

Definition 3.3.6

Als charakteristische Funktion der Zufallsvariablen X bezeichnen wir die Funktion

$$f(u) = E[\exp(juX)] = \int_{-\infty}^{\infty} \exp(jux) \cdot p(x) \, dx \quad (3.26)$$

Dabei ist u eine reelle Zahl und $j = \sqrt{-1}$ die imaginäre Einheit. ([26], S. 132)

Die rechte Seite der Gleichung (3.26) zeigt, daß man $f(u)$ auch als inverse Fouriertransformierte der Wahrscheinlichkeitsdichte $p(x)$ auffassen kann. Auf ähnliche Weise kann man auch k – dimensionale charakteristische Funktionen definieren, wobei k eine beliebig große, aber endliche natürliche Zahl ist.

Definition 3.3.7

Es sei $\vec{X} = (X_1, X_2, \dots, X_k)$ ein Zufallsvektor und $\vec{u} = (u_1, u_2, \dots, u_k)$ ein Vektor, dessen Elemente reelle Zahlen sind. Das Skalarprodukt von $\vec{X}$ und $\vec{u}$ ist

$$\vec{u}^T \vec{X} = u_1 X_1 + u_2 X_2 + \dots + u_k X_k. \quad (3.27)$$

$\vec{u}^T$ bedeutet dabei "u transponiert", mit

$$\vec{u}^T = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_k \end{bmatrix}.$$

Die charakteristische Funktion $f_X(\vec{u})$ des Vektors $\vec{X}$ ist dann definiert als

$$f_X(\vec{u}) = f(u_1, u_2, \dots, u_k) = E[\exp(j \vec{u}^T \vec{X})], \quad (3.28)$$

$$= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \{j(u_1 x_1 + u_2 x_2 + \dots + u_k x_k)\} p_X(x_1, x_2, \dots, x_k) dx_1 \dots dx_k$$

wenn die Wahrscheinlichkeitsdichte $p(\vec{x})$ stetig ist ([14], S.101).

Die charakteristische Funktion des Zufallsvektors $\vec{X}$ ist damit die multidimensionale Fouriertransformierte der Wahrscheinlichkeitsdichte $p(\vec{x})$.

Die große Bedeutung der charakteristischen Funktion in der Stochastik ergibt sich aus den folgenden beiden Eigenschaften. Sie ermöglichen es, die Berechnung der Verteilungsfunktion bzw. der Wahrscheinlichkeitsdichte auf eine Bestimmung der charakteristischen Funktion zurückzuführen. Durch dieses Verfahren lassen sich viele Probleme der Stochastik wesentlich leichter lösen als durch die direkte Methode mit Hilfe der Verteilungsfunktion bzw. der Wahrscheinlichkeitsdichte.

Der Eindeutigkeitssatz gilt sowohl für eindimensionale als auch für mehrdimensionale Verteilungen.

Satz 3.3.8

Zwei Verteilungsfunktionen $F_1(x)$ und $F_2(x)$ sind dann und nur dann identisch, wenn ihre charakteristischen Funktionen identisch sind. ([42], S. 28)

Der folgende Satz von Lévy ermöglicht es, die Verteilungsfunktionen aus der charakteristischen Funktion zu bestimmen. Da wir die charakteristische

Funktion als Funktion der Verteilungsfunktion definiert haben, folgt aus dem Satz von Lévy und dem Eindeutigkeitssatz, daß die Bestimmung der Verteilung und die Bestimmung der charakteristischen Funktion gleichwertig sind.

Satz 3.3.9

Ist $f(u)$ die charakteristische Funktion der Verteilungsfunktion $F(x)$, dann gilt

$$F(a+b) - F(a-b) = \lim_{T \rightarrow \infty} \frac{1}{\pi} \int_{-T}^T \frac{\sin(bu)}{u} \cdot \exp(-jua) \cdot f(u) \, du, \quad (3.29)$$

unter der Voraussetzung, daß $a + b$ und $a - b$ (mit $b > 0$) Stetigkeitsstellen der Verteilungsfunktion $F(x)$ sind ([26], S.144).

Unter bestimmten Bedingungen kann man auch die Wahrscheinlichkeitsdichte $p(x)$ direkt aus der charakteristischen Funktion bestimmen.

Satz 3.3.10

Ist die charakteristische Funktion im Intervall $(-\infty, \infty)$ absolut integrierbar, dann ist die zugehörige Wahrscheinlichkeitsdichte $p(x)$ stetig, und es gilt die Formel ([42], S.33)

$$p(x) = F'(x) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-jux) \cdot f(u) du \quad (3.30)$$

Die Formel (3.30) ist die Umkehrung von (3.26). Sie zeigt außerdem, daß die Dichte die Fouriertransformierte der charakteristischen Funktion ist.

Der Satz von Lévy kann auf k -dimensionale Verteilung und Wahrscheinlichkeitsdichten erweitert werden (Inversionssatz im R_k).

Satz 3.3.11

Ist $F(\vec{x})$ die Verteilungsfunktion eines Zufallsvektors $\vec{X}$ mit der charakteristischen Funktion $f_{\vec{X}}(\vec{u}) = f(u_1, u_2, \dots, u_k)$ und Δ ein k -dimensionales Intervall, das alle Punkte $(x_1, x_2, \dots, x_k)$ enthält, so daß $a_i - b \leq x_i < a_i + b$ ($i = 1, 2, \dots, k$), dann ist

$$F(\vec{a} + \vec{b}) - F(\vec{a} - \vec{b}) =$$

$$= \lim_{T \rightarrow \infty} \frac{1}{\pi^k} \int_{-T}^T \dots \int_{-T}^T \left[\prod_{i=1}^k \frac{\sin(b_i u_i)}{u_i} \right] \exp(-j\vec{a}\vec{u}) f(\vec{u}) d^k u \quad (3.31)$$

mit $\vec{a} = (a_1, a_2, \dots, a_k)$ und $d^k u = du_1 \cdot du_2 \cdot \dots \cdot du_k$; vorausgesetzt, daß Δ ein Stetigkeitsintervall von $F(\vec{x})$ ist. ([41] S.23, [26] S.152)

Die Bestimmung der k -dimensionalen Wahrscheinlichkeitsdichte ermöglicht der folgende Satz.

Satz 3.3.12

Ist die charakteristische Funktion $f(u_1, u_2, u_3, \dots, u_k)$ über den k -dimensionalen Raum R_k absolut integrierbar, dann ist die entsprechende Verteilungsfunktion stetig und die Wahrscheinlichkeitsdichte gegeben durch die Formel

$$\begin{aligned}
 & p(x_1, x_2, \dots, x_k) \\
 &= \frac{1}{(2\pi)^k} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp\{-j(u_1 x_1 + \dots + u_k x_k)\} f(u_1, u_2, \dots, u_k) du_1 \dots du_k
 \end{aligned}
 \tag{3.32}$$

([41] S.24) Die k -dimensionale Wahrscheinlichkeitsdichte des Zufallsvektors $\vec{X}$ ist damit die k -dimensionale Fouriertransformierte der entsprechenden charakteristischen Funktion.

3.4 Wahrscheinlichkeitsdichte nach der Filterung

Wie wir gesehen haben, bildet ein lineares Filter einen Zufallsvektor am Eingang auf eine Zufallsvariable am Ausgang ab. Durch die Einführung der charakteristischen Funktion sind damit zwei Wege zur Bestimmung der Dichte am Ausgang des Filters denkbar:

1. Die direkte Berechnung der Dichte aus der k -dimensionalen Wahrscheinlichkeitsdichte des Eingangsprozesses.
2. Bestimmung der eindimensionalen charakteristischen Funktion am Ausgang mit Hilfe der k -dimensionalen charakteristischen Funktion am Eingang und anschließende Berechnung der Wahrscheinlichkeitsdichte durch Fouriertransformation der eindimensionalen charakteristischen Funktion am Ausgang.

Es wird sich zeigen, daß die zweite Methode, die auf den ersten Blick aufwendiger wirkt, wesentlich einfacher durchzuführen ist als die erste. In manchen Fällen (siehe Kapitel 4.) ist die k -dimensionale Wahrscheinlichkeitsdichte des Eingangsprozesses im Gegensatz zur charakteristischen

Funktion sogar unbekannt, so daß das andere Verfahren überhaupt nicht in Frage kommt.

Zur numerischen Berechnung der Dichte benötigen wir in beiden Fällen unsere Näherung für den Ausgangsprozess

$$Y_i^k = \sum_{n=0}^{k-1} h(n) \cdot X(i-n) \quad (3.33)$$

Wir haben bereits festgestellt, daß der Eingangsprozeß zu jedem Zeitpunkt i ein Zufallsvektor ist. Für den Eingangsvektor vereinbaren wir die folgende Schreibweise

$$\vec{X}_i = (X(i-k+1), X(i-k+2), \dots, X(i-1), X(i)) \quad (3.34)$$

wobei i der Beobachtungszeitpunkt ist. $\vec{X}_i$ besteht also aus einer Zufallsvariablen der Gegenwart und $k-1$ Zufallsvariablen der Vergangenheit des Eingangsprozesses. Für die Wahrscheinlichkeitsdichte des Eingangsvektors $\vec{X}_i$ schreiben wir

$$p_i(\vec{x}) = p(x_{i-k+1}, x_{i-k+2}, \dots, x_{i-1}, x_i) \quad (3.35)$$

Die zugehörige charakteristische Funktion bzw. Verteilungsfunktion bezeichnen wir mit $f_i(\vec{u})$ und $F_i(\vec{x})$.

Diese Schreibweise ist deshalb sinnvoll, weil wir zunächst keine einschränkenden Annahmen über den Eingangsprozeß machen wollen. Bei nichtstationären Prozessen kann der Vektor $\vec{X}_i$ zu jedem Zeitpunkt i eine andere Wahrscheinlichkeitsdichte $p_i(\vec{x})$ besitzen. Im stationären Fall sind aber alle $p_i(\vec{x})$ identisch. Die Schätzung k -ter Ordnung des Eingangsprozesses kann man nun bequem in Vektorschreibweise formulieren, indem wir für die Impulsantwort

$$\vec{h} = (h(k-1), h(k-2), \dots, h(1), h(0)) \quad (3.36)$$

schreiben.

Die Näherung k -ter Ordnung für die Zufallsvariablen des Ausgangsprozesses $\{X_i, i \in \mathbb{Z}\}$ ist dann

$$Y_i^k = \vec{h}^T \cdot \vec{X}_i = \sum_{n=0}^{k-1} h(n) \cdot X(i-n) =$$

$$X(i-k+1) \cdot h(k-1) + X(i-k+2) \cdot h(k-2) + \dots + X(i) \cdot h(0) \quad (3.37)$$

Die rechte Seite der Gleichung legt nahe, daß man den Ausgangsprozeß als mit $\vec{h}$ gewichtetes Mittel des Eingangsprozesses ansehen kann. Die Gewichte $h(k-1), \dots, h(1), h(0)$ wandern dabei mit i über den Eingangsprozeß. Man spricht deshalb auch von einem "Prozeß gleitender Durchschnitte" (engl. moving average process). Dennoch handelt es sich bei dem hier untersuchten Prozeß am Ausgang des Filters nicht um einen Moving-Average-Prozeß im

klassischen Sinne (z.B. Box und Jenkins, [6]). Unter einem MA-Prozeß versteht man für gewöhnlich gefiltertes weißes Rauschen. Der Prozeß ist dann stationär.

Bei unserer Untersuchung wird aber Stationarität nicht vorausgesetzt.

3.5 Berechnung mit Hilfe der multidimensionalen Dichte

Wie wir noch sehen werden, erhält man einen besonders einfachen Ausdruck für die Wahrscheinlichkeitsdichte nach der Filterung, wenn man die multidimensionale charakteristische Funktion benutzt. (Kapitel 3.6).

Auch bei der Anwendung auf einen quadratischen Detektor (Kapitel 4.) wird ausschließlich die charakteristische Funktion verwandt.

Wir beschränken uns in diesem Kapitel daher auf folgendes Beispiel:

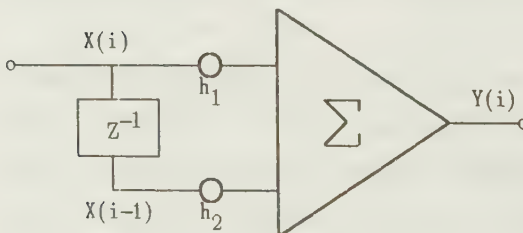


Bild 3.2 Ein nichtrekursives Filter

Bild 3.2 zeigt die Schaltungsstruktur eines nichtrekursiven Digitalfilters. Es besteht aus einem Speicher, der eine Verzögerung um jeweils ein Abtastintervall (Z^{-1}) realisiert, zwei Multiplizierern, die das Eingangssignal mit den Filterkoeffizienten $h(n)$ bewerten, und einem Addierer (Σ).

Die Impulsantwort $h(n)$ und die Filterkoeffizienten sind bei solchen Filtern identisch.

Wir nehmen außerdem an, daß der Eingangsprozeß stationär ist. Seine Verteilungsfunktion ist dann unabhängig vom Beobachtungszeitpunkt i .

Wegen der endlichen Länge $k = 2$ der Impulsantwort benötigen wir keine Näherung für die Zufallsvariable Y . Für die Wahrscheinlichkeitsdichte des zweidimensionalen Eingangsvektors schreiben wir

$$p(x_1, x_2) = F'(x_1, x_2) \quad (3.38)$$

und

$$F(X_1 < x_1, X_2 < x_2) \quad (3.39)$$

Zu einem festen Zeitpunkt $i = 2$ ist die Zufallsvariable am Ausgang gegeben durch

$$Y = \vec{h}^T \cdot \vec{X} = h_2 X_1 + h_1 X_2 \quad (3.40)$$

Faßt man die Realisationen (x_1, x_2) des Zufallsvektors $\vec{X}$ als Punkte auf einer Fläche auf, dann ist die Verteilungsfunktion $F(y)$ definiert durch die Wahrscheinlichkeit, daß der Punkt (x_1, x_2) unterhalb einer Geraden mit

$$y = h_2 x_1 + h_1 x_2 \quad (3.41)$$

liegt.

Die Verteilungsfunktion der Zufallsvariablen Y ist dann nach ihrer Definition

$$\begin{aligned} F(y) &= P(Y < y) = P(h_2 X_1 + h_1 X_2 < y) = \\ &= P(X_2 < \frac{1}{h_1} (y - h_2 x_1)) \end{aligned} \quad (3.42)$$

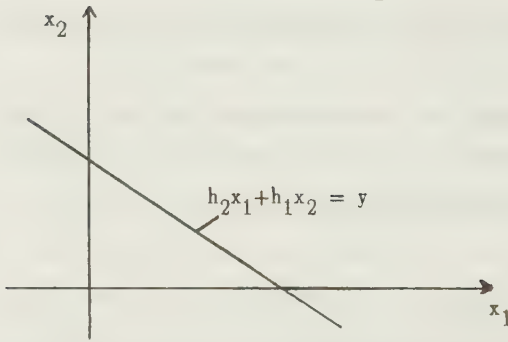


Bild 3.3 Bestimmung der Verteilungsfunktion nach der Filterung

Ist die zweidimensionale Verteilung $p(x_1, x_2)$ bekannt, so erhalten wir

$$F(y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\frac{1}{h_2} (y - h_1 x_2)} p(x_1, x_2) dx_1 dx_2 \quad (3.43)$$

Leiten wir nun nach y ab, erhalten wir die Wahrscheinlichkeitsdichte nach

der Filterung

$$p(y) = \int_{-\infty}^{\infty} p(x_1, x_2 = \frac{1}{h_1}(y - h_2 x_1)) dx_1 \quad (3.44)$$

Sind die Zufallsvariablen X_1 und X_2 statistisch unabhängig, so ist ihre zweidimensionale Wahrscheinlichkeitsdichte das Produkt der Wahrscheinlichkeitsdichten der einzelnen Zufallsvariablen:

$$p(x_1, x_2) = p(x_1) \cdot p(x_2) \quad (3.45)$$

Aus (3.44) wird dann ein Faltungsintegral.

$$p(y) = \int_{-\infty}^{\infty} p(x_1) \cdot p(x_2 = \frac{1}{h_1}(y - h_2 x_1)) dx_1 \quad (3.46)$$

Solche Überlegungen sind relativ bekannt, wurden aber bisher nicht auf die Filterung angewandt. Man findet sie z.B. bei Davenport und Root ([15], S.36), Fisz ([26], S. 81) und Schlitt ([54] S. 100)

Schlitt geht aber irrtümlich davon aus, daß Gleichung (3.44) in ein Faltungsintegral übergeht, wenn die Zufallsvariablen X_1 und X_2 unkorreliert sind.

Für die Produktform der Wahrscheinlichkeitsdichte (3.45) wird aber die statistische Unabhängigkeit der Zufallsvariablen benötigt.

Statistisch unabhängige Zufallsvariablen sind zwar stets unkorreliert, die Umkehrung gilt aber im allgemeinen nicht.

Zwei unkorrelierte Zufallsvariablen sind also nicht notwendig statistisch unabhängig. Ein Beispiel für unkorrelierte, aber abhängige Variablen findet man bei Cramér ([14], S.279).

Eine wichtige Ausnahme bilden hier allerdings die normalverteilten Zufallsvariablen: Unkorrelierte normalverteilte Zufallsvariablen sind immer statistisch unabhängig.

3.6 Charakteristische Funktion nach der Filterung

Mit dem Satz (3.3.10) haben wir eine eindeutige Verknüpfung zwischen der charakteristischen Funktion und der Wahrscheinlichkeitsdichte einer Zufallsvariablen. Wir können daher die Wahrscheinlichkeitsdichte $p_1(y)$ nach der Filterung bestimmen, wenn es gelingt, die eindimensionale charakteristische Funktion $f_1^k(u)$ der Zufallsvariablen

$$Y_i^k = \hat{h}^T \cdot \hat{X} = \sum_{n=0}^{k-1} h(n) \cdot X(i-n) \quad (3.47)$$

für beliebig große k zu ermitteln, wenn zusätzlich der Grenzwert

$$f_i(u) = \lim_{k \rightarrow \infty} f_i^k(u) \quad (3.48)$$

existiert und die Fouriertransformierte von $f_i(u)$ bekannt ist – oder berechnet werden kann mit

$$p_i(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-juy) \cdot f_i(u) du. \quad (3.49)$$

Die Bestimmung der eindimensionalen charakteristischen Funktion von Y_i^k ist im allgemeinen, d.h. wenn man keine einschränkenden Annahmen über den Eingangsprozeß $\{X_i, i \in \mathbb{Z}\}$

macht, nur dann möglich, wenn die k -dimensionale charakteristische Funktion bzw. Verteilungsfunktion von $\{X_i, i \in \mathbb{Z}\}$ bekannt ist. Man kann nämlich zeigen, daß die Kenntnis der Verteilungen aller Linearkombinationen $Y_i^k = \vec{h}^T \cdot \vec{X}$ der Elemente $(X_1, X_2, \dots, X_k)$ eines Zufallsvektors $\vec{X}$ gleichbedeutend ist mit der Kenntnis der k -dimensionalen Verteilung $F_i(\vec{x})$ bzw. Wahrscheinlichkeitsdichte $p_i(\vec{x})$ oder charakteristischen Funktion $f_i(\vec{u})$ des Zufallsvektors.

Diesen Zusammenhang haben Cramér und Wold [13] mit Hilfe der multidimensionalen charakteristischen Funktion bewiesen.

Satz 3.6.1 (Satz von Cramér und Wold)

Die k -dimensionale Wahrscheinlichkeitsdichte $p(\vec{x})$ eines Zufallsvektors $\vec{X} = (X_1, X_2, \dots, X_k)$ ist eindeutig bestimmt durch die Klasse aller eindimensionalen Wahrscheinlichkeitsdichten der Zufallsvariablen

$$Y = S(\vec{X}) = h_1 X_1 + h_2 X_2 + \dots + h_k X_k \quad (3.50)$$

wobei alle h_i die Menge der reellen Zahlen durchlaufen.

Beweis.

Alle Elemente von $\vec{X} = (X_1, X_2, \dots, X_k)$ sind Zufallsvariablen. Damit ist auch $Y = h_1 X_1 + h_2 X_2 + \dots + h_k X_k$ eine Zufallsvariable, solange

$$\vec{h} \neq (0, 0, \dots, 0) \quad (3.51)$$

Die eindimensionale charakteristische Funktion von Y ist:

$$\begin{aligned} f_Y(u) &= E[\exp(juY)] = E[\exp(juS(\vec{X}))] = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \{ju(h_1 x_1 + h_2 x_2 + \dots + h_k x_k)\} p_X(x_1, x_2, \dots, x_k) dx_1 \dots dx_k \end{aligned} \quad (3.52)$$

Die k -dimensionale charakteristische Funktion des Zufallsvektors $\vec{X}$ ist dagegen nach Definition (3.36)

$$f_{\vec{X}}(\vec{v}) = f(v_1, v_2, \dots, v_k) = E[\exp(j\vec{v}^T \cdot \vec{X})]$$

$$= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \{j(v_1 x_1 + v_2 x_2 + \dots + v_k x_k)\} p_{\vec{X}}(x_1, x_2, \dots, x_k) dx_1 \dots dx_k$$
(3.53)

Ein Vergleich von (3.52) und (3.53) zeigt, daß ein direkter Zusammenhang zwischen der k -dimensionalen charakteristischen Funktion $f_{\vec{X}}(\vec{v})$ und der eindimensionalen charakteristischen Funktion $f_{\vec{Y}}(u)$ besteht, mit

$$f_{\vec{Y}}(u) = f_{\vec{X}}(u \cdot \vec{h}). \quad (3.54)$$

Für $u = 1$ ist demnach die eindimensionale charakteristische Funktion $f_{\vec{Y}}(u)$ identisch mit der k -dimensionalen charakteristischen Funktion $f_{\vec{X}}(\vec{v})$ im Punkt $\vec{v} = \vec{h}$.

Dies gilt offenbar auch für $\vec{h} = (0, 0, \dots, 0)$. Nehmen alle h_i unabhängig voneinander jeden Wert auf der reellen Achse an, dann ist $f_{\vec{X}}(\vec{v})$ vollständig determiniert. Wegen der Gültigkeit des Eindeutigkeitssatzes ist der Satz von Cramér und Wold damit bewiesen.

Interessanterweise beinhaltet der Beweis dieses Satzes auch den Beweis für seine Umkehrung, eine Eigenschaft, die unbekannt zu sein scheint.

Die in dieser Arbeit erstmals auf nachrichtentechnische Probleme angewandte Umkehrung ermöglicht es, eine Lösung für die charakteristische Funktion nach der Filterung explizit anzugeben, wenn die multidimensionale charakteristische Funktion vor der Filterung bekannt ist. Die Umkehrung des Satzes von Cramér und Wold kann man wie folgt formulieren:

Satz 3.6.2

Ist $f_{\vec{X}}(\vec{v})$ die k -dimensionale charakteristische Funktion des Zufallsvektors $\vec{X} = (X_1, X_2, \dots, X_k)$ und die Zufallsvariable $Y = \vec{h}^T \cdot \vec{X}$ eine Linearkombination seiner Elemente, dann ist die eindimensionale charakteristische Funktion $f_Y(u)$ gegeben durch $f_Y(u) = f_{\vec{X}}(u \cdot \vec{h})$.

Beweis. Siehe Satz 3.6.1 Gleichung (3.54)

Anwendungen für den zuletzt genannten Satz sind vermutlich deshalb sehr selten, weil in vielen Fällen die Bestimmung der multidimensionalen charakteristischen Funktion ein schwieriges Problem darstellt.

Für die Berechnung der Wahrscheinlichkeitsdichte nach der Filterung ist die Bestimmung der multidimensionalen Verteilung bzw. charakteristischen Funktion aber nicht zu vermeiden, wie der Satz von Cramér und Wold zeigt.

Einfachere (z.B. eindimensionale) statistische Kenngrößen determinieren die eindimensionale Dichte nach der Filterung nur dann, wenn $\vec{X}$ eine spezielle Wahrscheinlichkeitsdichte hat und diese bereits durch einfachere Kenngrößen vollständig beschrieben werden kann. Das bekannteste Beispiel ist die k-dimensionale Normalverteilung mit der charakteristischen Funktion

$$f(\vec{u}) = \exp\{\vec{j} \cdot \vec{b} \cdot \vec{u}^T - \frac{1}{2} \vec{u} \cdot \vec{\Lambda} \cdot \vec{u}^T\} \quad (3.55)$$

Diese ist bereits durch die Autokovarianzmatrix $\vec{\Lambda}$ und den Mittelwertvektor $\vec{b}$ eindeutig festgelegt. Die eindimensionale Wahrscheinlichkeitsdichte kann daher bestimmt werden, wenn $\vec{\Lambda}$ und $\vec{b}$ bekannt sind.

Die wichtigste Eigenschaft von Satz 3.6.2 ist aber, daß man ihn ohne Einschränkung benutzen kann; er gilt für stationäre und nichtstationäre Prozesse und solche mit beliebigen statistischen Abhängigkeiten zwischen den Zufallsvariablen, da er ohne solche Annahmen für den Eingangsprozeß hergeleitet wurde. Es ist damit also möglich, ein universelles Verfahren zur Bestimmung der Dichte nach der linearen Filterung anzugeben, wobei allerdings die Kenntnis der multidimensionalen charakteristischen Funktion des Eingangsprozesses vorausgesetzt wird.

Den Zusammenhang zwischen der charakteristischen Funktion des Zufallsvektors $\vec{X}$ und der charakteristischen Funktion der Zufallsvariablen $Y = \vec{h}^T \cdot \vec{X}$ nach der Filterung kann man sich mit Satz (3.6.2) graphisch sehr

gut veranschaulichen. Die eindimensionale charakteristische Funktion $f_y(u)$ einer Linearkombination der Elemente eines Zufallsvektors $\vec{X}$ bzw. nach einem Filter mit der Impulsantwort $\vec{h}$ entspricht der k -dimensionalen charakteristischen Funktion $f_x(\vec{v})$ des Zufallsvektors $\vec{X}$ entlang der Geraden $\vec{v} = u \cdot \vec{h}$, wobei k die Anzahl der Gewichte $h(n)$ bzw. die Länge der Impulsantwort ist.

Bild 3.4 zeigt die zweidimensionale charakteristische Funktion der zentralen χ^2 -Verteilung, (Kapitel 4.2, GL. 4.24, $\lambda_{1,1} = \lambda_{2,2} = 0,1$ $\lambda_{1,2} = \lambda_{2,1} = 0$) vor der Filterung und die eindimensionale charakteristische Funktion nach einem Filter mit der Impulsantwort

$$h_2 = 1 \text{ und } h_1 = 1.$$

In Bild 3.5 haben wir die gleiche zweidimensionale charakteristische Funktion wie in Bild 3.4, die Impulsantwort des Filters ist aber $h_1 = 0,25$ und $h_2 = 1$.

Der Satz von Cramér und Wold kann ebenfalls geometrisch veranschaulicht werden. (ebenfalls Bild 3.4 und 3.5). Man könnte ihn dann folgendermaßen formulieren:

Die k -dimensionale charakteristische Funktion $f_x(\vec{v})$ ist eindeutig festgelegt durch die Funktionswerte entlang aller Geraden $\vec{v} = u \cdot \vec{h}$, die den Koordinatenursprung schneiden, diese Funktionswerte entsprechen dann den eindimensionalen charakteristischen Funktionen der Linearkombinationen $Y = \vec{h}^T \cdot \vec{X}$.

Bei der Anwendung dieser Sätze auf die Berechnung der Wahrscheinlichkeitsdichte nach der Filterung tritt allerdings das Problem auf, daß die Impulsantwort $\vec{h}$ im allgemeinen unendlich lang ist; die charakteristische Funktion ist aber nur für endlich viele Dimensionen definiert. Ist die charakteristische Funktion des Eingangsprozesses für beliebig viele Dimensionen bekannt, kann dieses Problem durch Bildung eines Grenzwertes gelöst werden.

Die eindimensionale charakteristische Funktion der Näherung k -ter Ordnung des Eingangsprozesses erhalten wir durch Abbrechen der Impulsantwort $\vec{h}$ im Punkt k und durch Anwendung des Satzes (3.6.2).

Die Näherung für Y_i nach der Filterung ist:

$$Y_i^k = \vec{h}^T \cdot \vec{X} = \sum_{n=0}^{k-1} h(n) \cdot X(i-n) \quad (3.56)$$

mit

$$\vec{h} = (h(k-1), \dots, h(1), h(0)) \quad (3.57)$$

Ist nun die k -dimensionale charakteristische Funktion $f_i(\vec{v})$ vor der Filterung bekannt (diese kann von Beobachtungszeitpunkt i abhängen), dann ist die eindimensionale charakteristische Funktion von Y_i^k mit Satz 3.6.2

$$f_i^k(u) = f_i(\vec{v} = u \cdot \vec{h}) = f_i(u \cdot h(k-1), \dots, u \cdot h(1), u \cdot h(0)). \quad (3.58)$$

Die Wahrscheinlichkeitsdichte der Zufallsvariablen Y_i^k ist nach Satz (3.3.10) die Fouriertransformierte der charakteristischen Funktion:

$$p_i^k(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-juy) \cdot f_i^k(u) du \quad (3.59)$$

Unter der Bedingung, daß der Grenzwert für die charakteristische Funktion von Y existiert mit

$$f_i(u) = \lim_{k \rightarrow \infty} f_i^k(u), \quad (3.60)$$

existiert auch der Grenzwert für die Wahrscheinlichkeitsdichte. Die Lösung für die eindimensionale Wahrscheinlichkeitsdichte nach linearer Filterung ist damit:

$$p_i(y) = \lim_{k \rightarrow \infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-juy) \cdot f_{x_i}(u \cdot h(k-1), \dots, u \cdot h(1), u \cdot h(0)) du \quad (3.61)$$

wobei $h(n)$ die Impulsantwort des Filters und f_{x_i} die k - dimensionale charakteristische Funktion des Eingangsvektors

$\vec{X}_i = (X(i-k+1), X(i-k+2), \dots, X(i-1), X(i))$ zum Zeitpunkt i ist.

Die Existenz des Grenzwertes in (3.6.1) folgt unter der Bedingung, daß der Grenzwert für die charakteristische Funktion existiert, aus dem Satz von Lévy und Cramér, der in der englischsprachigen Literatur auch als "Continuity theorem" bekannt ist.

Satz 3.6.3

Konvergiert eine Folge von charakteristischen Funktionen $f^k(u)$ an jeder Stelle u ($-\infty < u < \infty$) gegen die Funktion $f(u)$, die in einem gewissen Intervall $|u| < a$ stetig ist, dann konvergiert die Folge der zugehörigen Verteilungsfunktionen $F^k(x)$ gegen die Verteilungsfunktion $F(x)$, deren charakteristische Funktion $f(u)$ ist.

Bei praktischen Anwendungen kann die Bestimmung des Grenzwertes für die charakteristische Funktion jedoch große Probleme bereiten. Man erhält dann aber zumindest eine Näherungslösung mit Gleichung (3.59).

Die Berechnung des Fourierintegrals in Gleichung (3.59) ist ebenfalls nicht immer möglich. Eine numerische Berechnung mittels diskreter Fouriertransformation führt hier ebenfalls zu einer Näherungslösung. Das wichtigste Problem ist damit die Bestimmung der multidimensionalen charakteristischen Funktion vor der Filterung.

Diese enthält dann bereits alle eindimensionalen charakteristischen Funktionen einer Näherung k -ter Ordnung, wie Satz 3.6.2 zeigt.

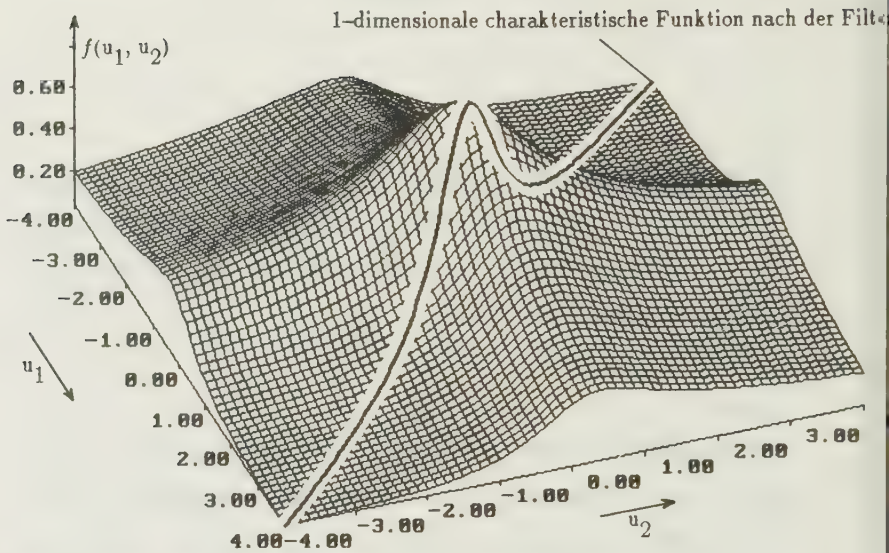


Bild 3.4a Realteil einer 2-dimensionalen charakteristischen Funktion vor der Filterung.

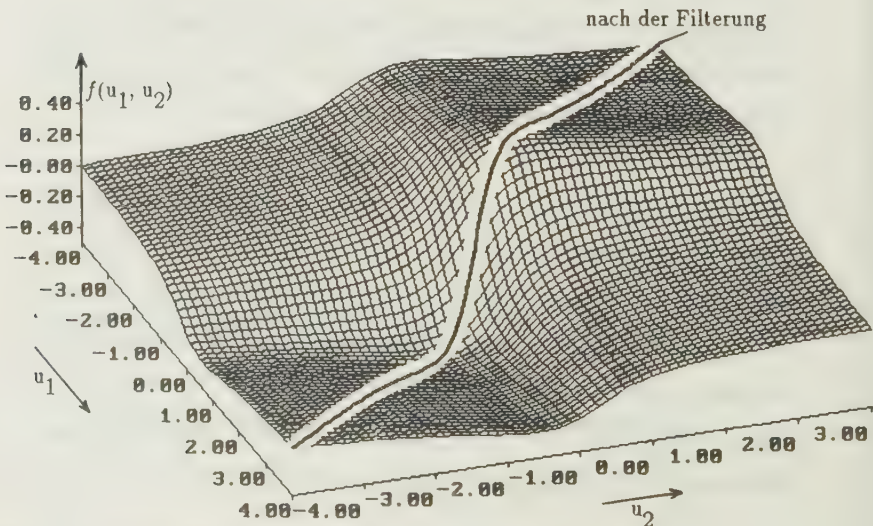


Bild 3.4b Imaginärteil einer 2-dimensionalen charakteristischen Funktion vor der Filterung.

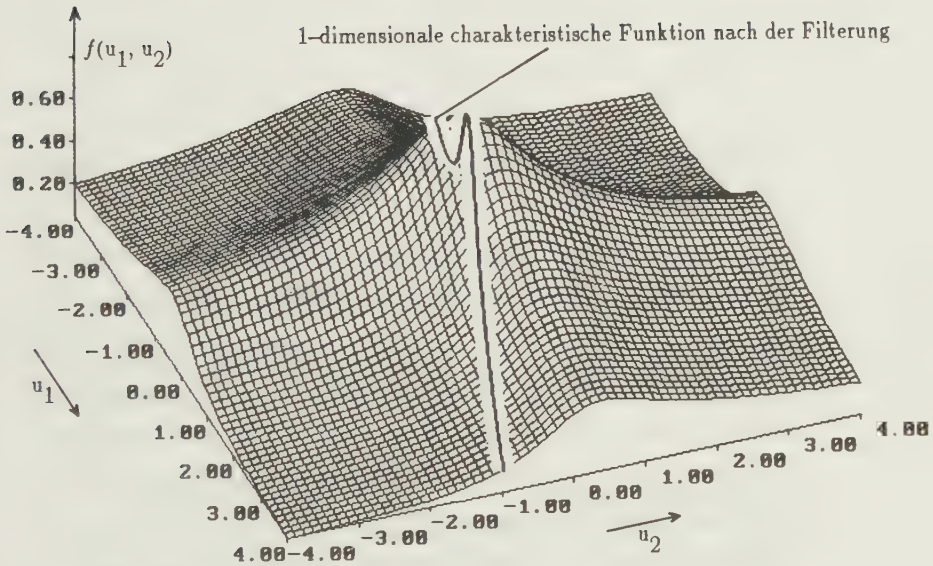


Bild 3.5a Realteil einer 2-dimensionalen charakteristischen Funktion vor der Filterung.

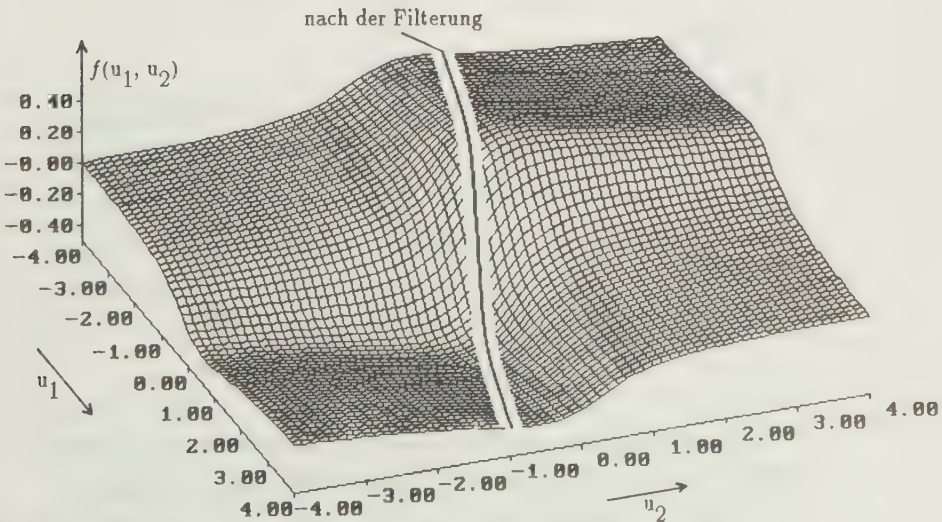


Bild 3.5b Imaginärteil einer 2-dimensionalen charakteristischen Funktion vor der Filterung.

4. Anwendung auf einen quadratischen Detektor

Wir wollen die vorangegangenen theoretischen Überlegungen nun auf eine in der Praxis häufig vorkommende Empfänger-Schaltung anwenden.

Diese besteht aus einem ZF-Bandpaß mit der Impulsantwort $g(n)$, einem Quadrierer und einem abschließenden NF-Tiefpaß mit der Impulsantwort $h(n)$. (Bild 4.1). Eingangsfolge ist die Summe aus einem deterministischen Signal $S(i)$ und weißen Rauschen $N(i)$. Alle vor dem ZF-Filter liegenden Bauelemente wie z.B. Mischer werden bei unseren Überlegungen nicht berücksichtigt, d.h. es wird angenommen, daß die Rauschleistung, die direkt in das ZF-Filter eingespeist wird, groß ist gegenüber allen anderen Rauschanteilen.

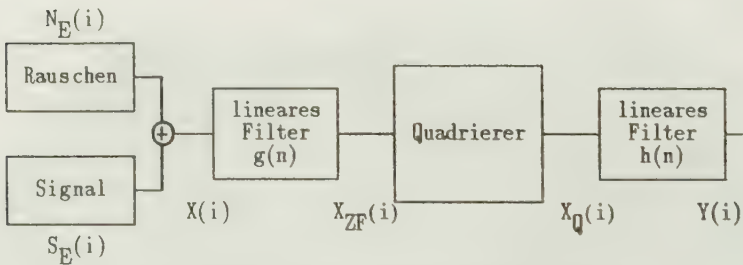


Bild 4.1 Ein diskreter quadratischer Detektor

Diese Anordnung ist aus mehreren Gründen gut geeignet, die vorgeschlagene Methode zur Berechnung der Verteilung nach linearer Filterung zu

demonstrieren:

1. Die Annahme, daß das Eingangssignal eine Überlagerung aus Signal und weißem (thermischem) Rauschen ist, ermöglicht es, die multidimensionale Wahrscheinlichkeitsdichte geschlossen zu bestimmen.
2. Nach der Quadrierstufe haben wir einen statistisch abhängigen, nicht normalverteilten Zufallsprozeß, dessen charakteristische Funktion ebenfalls in geschlossener Form bestimmt werden kann.
3. Das abschließende lineare (NF-) Filter $h(n)$ verändert die Wahrscheinlichkeitsdichte erneut; diese kann mit Hilfe der in Kapitel 3.4 vorgeschlagenen Methode berechnet werden, und zwar für beliebige Filter.
4. Dieses Problem ist für analoge Filter bereits untersucht worden (Emerson [25]).

4.1 Charakteristische Funktion nach der ZF-Filterung

Betrachten wir zunächst die Wahrscheinlichkeitsdichte nach dem ersten linearen Filter $g(n)$. Das Eingangssignal ist die Summe eines deterministischen Signals $S_E(i)$ und Rauschen $N_E(i)$

$$X(i) = S_E(i) + N_E(i) \quad (4.1)$$

dabei ist $S(i)$ ein beliebiges bandbegrenztes Signal mit der oberen Grenzfrequenz $f_g \leq \frac{1}{2 \cdot T}$. Es liegt in abgetasteter Form vor.

Die Abtastfrequenz ist dabei $f_a = \frac{1}{T}$. $N(i)$ sei durch Abtasten von thermischem Rauschen entstanden.

Das Filter $g(n)$ kann als Faltungsoperator aufgefaßt werden. Wegen der Linearität dieses Operators kann man den gefilterten deterministischen Anteil

$$S_{ZF}(i) = \sum_{n=-\infty}^{\infty} g(n) \cdot S_E(i-n) = g(i) * S_E(i) \quad (4.2)$$

und den gefilterten stochastischen Anteil

$$N_{ZF}(i) = \sum_{n=-\infty}^{\infty} g(n) \cdot N_E(i-n) = g(i) * N_E(i) \quad (4.3)$$

getrennt betrachten (der $*$ bedeutet "gefaltet mit"), denn es gilt

$$\begin{aligned} X_{ZF}(i) &= \sum_{n=-\infty}^{\infty} g(n) \cdot [S_E(i-n) + N_E(i-n)] = \\ &= \sum_{n=-\infty}^{\infty} g(n) \cdot S_E(i-n) + \sum_{n=-\infty}^{\infty} g(n) \cdot N_E(i-n) = S_{ZF}(i) + N_{ZF}(i). \end{aligned} \quad (4.4)$$

Die Wahrscheinlichkeitsdichte und charakteristische Funktion des stochastischen Anteils können bestimmt werden, wenn man voraussetzt, daß es sich bei $N_E(i)$ um thermisches Rauschen handelt, und wenn die Impulsantwort $g(n)$ bekannt ist.

Der Prozeß am Eingang des Filters ist – konstante Temperatur vorausgesetzt – stationär, mittelwertfrei, hat eine multidimensionale

Normalverteilung und ein bis 10^{13} Hz näherungsweise konstantes Leistungsdichtespektrum. ([48] S.482, [64]). Am Ausgang des Filters $g(n)$ erhalten wir ebenfalls eine Normalverteilung, da jede lineare Transformation eines normalverteilten Prozesses selbst normalverteilt ist ([14] S. 313).

Die Wahrscheinlichkeitsdichte der multidimensionalen Normalverteilung ist

$$p_N(\vec{z}) = \frac{|\hat{\Lambda}|^{1/2}}{(2\pi)^{k/2}} \exp\left\{-\frac{1}{2} \cdot \vec{z} \cdot \hat{\Lambda} \cdot \vec{z}^T\right\} \quad (4.5)$$

wobei $\hat{\Lambda} = \hat{\Lambda}^{-1}$ die Inverse der Kovarianzmatrix $\hat{\Lambda}$ und $|\hat{\Lambda}|$ die entsprechende Determinante ist.

Die multidimensionale charakteristische Funktion von N_{ZF} lautet dann ([41] S.28):

$$\begin{aligned} f_N(\vec{u}) &= f(u_1, u_2, \dots, u_k) \\ &= \frac{|\hat{\Lambda}|^{1/2}}{(2\pi)^{k/2}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp\left\{j \cdot \vec{u} \cdot \vec{z} - \frac{1}{2} \cdot \vec{z} \cdot \hat{\Lambda} \cdot \vec{z}^T\right\} d^k z = \\ &= \exp\left\{-\frac{1}{2} \cdot \vec{u} \cdot \hat{\Lambda} \cdot \vec{u}^T\right\}. \end{aligned} \quad (4.6)$$

$d^k z$ ist dabei ein k -dimensionales Volumenelement.

Die Wahrscheinlichkeitsdichte und charakteristische Funktion von mittelwertfreien normalverteilten Prozessen ist also durch die Autokovarianzmatrix $\hat{\Lambda}$ vollständig beschrieben. Da der Mittelwert gleich

Null ist, entspricht die Autokovarianz der Autokorrelation des Prozesses.

Bei stationären Prozessen sind außerdem die Werte der Elemente $\lambda_{i,j}$ der Autokovarianzmatrix $\vec{\lambda}$ nur von der Differenz $\ell = i-j$ abhängig, mit ([18] S.596)

$$\lambda_{i,j} = R(i-j), \quad (4.7)$$

Wir benötigen also nur die Autokorrelationsfunktion $R(\ell)$ von $N_{ZF}(i)$, um die multidimensionale charakteristische Funktion $f_{ZF}^*(u)$ bestimmen zu können.

Nach Middleton ([48], S. 162) erhalten wir die Autokorrelationsfunktion von gefiltertem weißen Rauschen direkt aus der Impulsantwort, indem wir $g(\ell)$ mit $g(-\ell)$ falten. Für den diskreten Fall ist

$$R(\ell) = \sum_{n=-\infty}^{\infty} g(n) \cdot g(\ell + n) = [g(\ell) * g(-\ell)] \quad (4.8)$$

Gleichung (4.8) bezeichnet man als deterministische Autokorrelationsfunktion des Filters $g(n)$.

Hatte der Prozeß vor der Filterung die Varianz σ^2 , dann ist die Autokorrelationsfunktion nach der Filterung

$$R_N(\ell) = \sigma^2 \cdot [g(\ell) * g(-\ell)] \quad (4.9)$$

Zur Bestimmung der k-dimensionalen charakteristischen Funktion der

Summe von Signal und Rauschen benutzen wir die Vektorschreibweise.

Das (deterministische) Signal ist dann

$$\vec{S}_{ZF} = (S(i-k+1), S(i-k+2), \dots, S(i-1), S(i)) \quad (4.10)$$

und der stochastische Anteil lautet

$$\vec{N}_{ZF} = (N(i-k+1), N(i-k+2), \dots, N(i-1), N(i)) \quad (4.11)$$

Dabei ist i der Beobachtungszeitpunkt. Der Zufallsvektor $\vec{X}_{ZF}$ am Ausgang des Detektors ist die Summe von Signal und Rauschen

$$\vec{X}_{ZF} = \vec{S}_{ZF} + \vec{N}_{ZF} \quad (4.12)$$

Der resultierende Vektor $\vec{X}_{ZF}$ ist ebenfalls ein Zufallsvektor.

Seine charakteristische Funktion $f_{ZF}(\vec{u})$ ist:

$$\begin{aligned} f_{ZF}(\vec{u}) &= E[\exp\{j \cdot \vec{u}^T \cdot \vec{X}_{ZF}\}] = E[\exp\{j \cdot \vec{u}^T \cdot (\vec{S}_{ZF} + \vec{N}_{ZF})\}] = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp\{j \cdot \vec{u}^T \cdot (\vec{z} + \vec{S}_{ZF})\} p_N(z_1, z_2, \dots, z_k) dz_1 \dots dz_k \\ &= \exp\{j \cdot \vec{u}^T \cdot \vec{S}_{ZF}\} \cdot \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp\{j \cdot \vec{u}^T \cdot \vec{z}\} p_N(z_1, z_2, \dots, z_k) dz_1 \dots dz_k \quad (4.13) \end{aligned}$$

Es gilt also

$$f_{ZF}(\vec{u}) = \exp\{\vec{j} \cdot \vec{u}^T \cdot \vec{\dot{S}}_{ZF}\} \cdot f_N(\vec{u}) \quad (4.14)$$

Mit Gleichung (4.6) und (4.14) haben wir für den Zufallsvektor $\vec{\dot{X}}_{ZF}$ am Eingang des Quadrierers die folgende charakteristische Funktion:

$$f_{ZF}(\vec{u}) = \exp\{\vec{j} \cdot \vec{u}^T \cdot \vec{\dot{S}}_{ZF} - \frac{1}{2} \sigma^2 \cdot \vec{u} \cdot \vec{\dot{\Lambda}} \cdot \vec{u}^T\} \quad (4.15)$$

mit

$$S_{ZF}(i) = g(i) * S_E(i), \quad i = 1, 2, \dots, k \quad (4.16)$$

$$\lambda_{i,j} = R_N(\ell = i-j) \quad i, j = 1, 2, \dots, k \quad (4.17)$$

und

$$R_N(\ell) = \sigma^2 \cdot [g(\ell) * g(-\ell)]. \quad (4.18)$$

dies entspricht der charakteristischen Funktion einer k -dimensionalen Normalverteilung mit der Kovarianzmatrix $\vec{\dot{\Lambda}}$ und dem Mittelwertvektor $\vec{\dot{S}}_{ZF}$.

4.2 Der Quadrierer

Wir suchen nun die charakteristische Funktion der Summe von Signal und Rauschen nach einer idealen Quadratur-Stufe. Ideal bedeutet hier zum einen, daß es sich um ein Bauelement mit einer perfekt quadratischen Kennlinie der Form $y = x^2$ handelt, zum anderen, daß es sich um eine gedächtnisfreie Nichtlinearität handelt, das heißt, daß das Bauelement nicht in der Lage ist,

Energie zu speichern. Außerdem darf das Rauschen des Quadrierers nur einen unwesentlichen Beitrag zur Gesamtrauschleistung des Systems liefern.

Wir haben nach der ZF-Filterung einen k -dimensional normalverteilten Vektor, der aus k Abtastwerten gebildet wurde mit

$$\vec{x}_{ZF} = (x_{ZF}(i-k+1), x_{ZF}(i-k+2), \dots, x_{ZF}(i-1), x_{ZF}(i)) \quad (4.19)$$

und suchen die Wahrscheinlichkeitsdichte des Vektors

$$\vec{x}_Q = (x_{ZF}(i-k+1)^2, x_{ZF}(i-k+2)^2, \dots, x_{ZF}(i-1)^2, x_{ZF}(i)^2) \quad (4.20)$$

Die Verteilung eines solchen Vektors $\vec{x}_Q$ heißt k -dimensionale χ^2 -Verteilung mit einem Freiheitsgrad, wenn der Vektor $\vec{x}_{ZF}$ (vor der Quadratur) mittelwertfrei war.

Enthält der Vektor $\vec{s}_{ZF}$ in (4.15) Elemente, die von Null verschieden sind, d.h. hat der Zufallsvektor $\vec{x}_{ZF}$ eine Normalverteilung mit einem Mittelwert, so bezeichnet man die Verteilung von $\vec{x}_Q$ als nichtzentrale k -dimensionale χ^2 -Verteilung mit einem Freiheitsgrad.

Bildet man die Summe aus n identisch verteilten, stochastisch unabhängigen Vektoren $\vec{x}_Q$ mit

$$\vec{y}_n = \sum_{i=1}^n \vec{x}_Q \quad (4.21)$$

so hat der Summenvektor $\vec{y}_n$ eine χ^2 -Verteilung mit n Freiheitsgraden.

Diese Verteilung beschreibt die Verteilung der Varianz einer Stichprobe aus n

normalverteilten Zufallsvektoren $\vec{X}$ und ist deshalb vor allem für die Statistik von Interesse. Die charakteristische Funktion von Y_n wurde erstmals von D.R. Jensen berechnet [32], dessen Untersuchung übrigens vom "National Institute of General Medical Sciences" gefördert wurde. Auf die Berechnung der Wahrscheinlichkeitsdichte von Signal und Rauschen nach einem Quadratischen Detektor wurden diese Ergebnisse bisher aber noch nicht angewandt. Jensen hat bewiesen, daß die charakteristische Funktion der Verteilung von Y_k wie folgt lautet

$$f_{Y_n}(\vec{u}) = |\vec{I} - 2 j \vec{U}_D \vec{\Lambda}|^{-n/2} \cdot \exp \{ j n \vec{b}^T (\vec{I} - 2 j \vec{U}_D \vec{\Lambda})^{-1} \vec{U}_D \vec{b} \} \quad (4.22)$$

dabei ist $\vec{U}_D$ eine Diagonalmatrix mit den Elementen $\{u_1, u_2, \dots, u_k\}$; $\vec{I}$ die Einheitsmatrix der Ordnung k und n die Anzahl der Freiheitsgrade. Für die von uns benötigte Verteilung des Quadrats eines normalverteilten Vektors mit der Kovarianzmatrix $\vec{\Lambda}$ und dem Mittelwertvektor $\vec{S}_{ZF}$ entspricht der Verteilung von $\vec{Y}_n$ für einen Freiheitsgrad, d.h. es gilt:

$$\begin{aligned} f_Q(\vec{u}) &= f_Q(u_1, u_2, \dots, u_k) \\ &= \frac{|\vec{\Lambda}|^{1/2}}{(2\pi)^{k/2}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \{ j \cdot \vec{u}^T (z_1^2, z_2^2, \dots, z_k^2) - \frac{1}{2} (\vec{z} - \vec{S}_{ZF}) \vec{\Lambda} (\vec{z} - \vec{S}_{ZF})^T \} d^k z \\ &= |\vec{I} - 2 j \vec{U}_D \vec{\Lambda}|^{-1/2} \cdot \exp \{ j \vec{S}_{ZF}^T (\vec{I} - 2 j \vec{U}_D \vec{\Lambda})^{-1} \vec{U}_D \vec{S}_{ZF} \} \quad (4.23) \end{aligned}$$

Wir haben damit eine geschlossene Lösung für die k - dimensionale

charakteristische Funktion von $\hat{x}_Q$, d.h. der Summe von Signal und Rauschen nach Filterung und Quadratur.

Ist der Prozeß vor der Quadratur mittelwertfrei, was in unserem Fall bedeutet, daß ein Signal fehlt und nur Rauschen in den Detektor eingespeist wird, so hat $\hat{x}_Q$ eine zentrale χ^2 - Verteilung mit der charakteristischen Funktion

$$f_Q(\vec{u}) = f_Q(u_1, u_2, \dots, u_k) = |\vec{I} - 2 \sum \vec{u}_D \vec{A}|^{-1/2} \quad (4.24)$$

Die Bestimmung der Wahrscheinlichkeitsdichteverteilung der multidimensionalen χ^2 - Verteilung ist in geschlossener Form – auch für den zentralen Fall und einen Freiheitsgrad – ein noch ungelöstes Problem. Selbst Reihenentwicklungen für den zweidimensionalen Fall stellen ein schweres Problem dar ([32], S. 812).

Bei Jensen finden wir aber eine geschlossene Lösung für den Mittelwert $\vec{m} = (m_1, m_2, \dots, m_k)$ und die Kovarianzmatrix $\vec{K}$ mit den Elementen $k_{i,j}$ nach der Quadratur; und zwar

$$m_i = \lambda_{i,i} + s_i^2 \quad (4.25)$$

und

$$k_{i,j} = \lambda_{i,j}^2 + 2 \cdot s_i \cdot s_j \cdot \lambda_{i,j}, \quad (4.26)$$

wobei s_i die Elemente des Vektors $\vec{s}_{ZF}$ sind.

Sind diese groß gegenüber λ_{ii} , was für unsere Anwendung gleichbedeutend

mit hohen Signal – Rausch – Abstand ist, so kann man mit folgender Näherung arbeiten, die ebenfalls von Jensen [32] stammt:

Satz 4.2.1

Wachsen die Elemente s_i über alle Schranken und bleibt die Kovarianzmatrix $\hat{\Lambda}$ dabei konstant, so nähert sich die Verteilung von $\hat{X}_Q$ mit immer kleiner werdendem Fehler einer Normalverteilung mit dem Mittelwert $\hat{m}$ und der Kovarianzmatrix $\hat{K}$.

Für unsere Anwendung bedeutet dieser Satz, daß wir für hohe Signal-Rausch-Abstände auch am Ausgang des Detektors (also nach dem NF-Filter $h(n)$) näherungsweise eine Normalverteilung haben.

Die Filterung mit $h(n)$ stellt nämlich eine lineare Transformation dar, die nur die Korrelation und den Mittelwert des Prozesses verändert.

Für geringe Signal-Rausch-Abstände, d.h. Elemente von $\hat{S}_{ZF}$, die in der Größenordnung der Elemente von $\hat{\Lambda}$ liegen, ist für die Wahrscheinlichkeitsdichte $p_Q(\hat{x})$ eine Näherungslösung mit Hilfe der k -dimensionalen diskreten Fourier-Transformation möglich; ein Weg, der aber offenbar noch nie beschritten wurde.

Aus Gründen der graphischen Darstellbarkeit, des Rechenzeit- und Speicherplatzbedarfs beschränken wir uns auf $k = 2$ Dimensionen; das Verfahren kann aber ohne weitere Probleme auf beliebig viele Dimensionen erweitert werden.

Bevor wir die diskrete Fouriertransformation anwenden können, muß die kontinuierliche charakteristische Funktion noch zweidimensional abgetastet

und gefenstert werden. Außerdem ist eine Rotation des Koordinatenursprungs notwendig.

Um sicherzustellen, daß wir als Ergebnis eine Wahrscheinlichkeitsdichte $p_Q(x_1, x_2)$ erhalten, die in einem vernünftigen Bereich von x_1 und x_2 liegt, und um dafür zu sorgen, daß die charakteristische Funktion mit einer ausreichend hohen Abtastfrequenz diskretisiert wird, führen wir eine Skalierungszahl $ISK \in \mathbb{N}$ ein. Als Skalierungsfaktor bezeichnen wir dann die Zahl

$$SK = 2^{ISK} \quad (4.27)$$

Wir tasten nun die charakteristische Funktion im Bereich von $-SK/2$ bis $SK/2$ ab und zwar mit einer Abtastrate von

$$T = 1/SK \quad (4.28)$$

und erhalten

$$N = (2^{ISK})^2 = 2^{2 \cdot ISK} = SK^2 \quad (4.29)$$

Abtastwerte.

Ist $f(u_1, u_2)$ die kontinuierliche charakteristische Funktion, wobei u_1 und u_2 Kreisfrequenzen sind, so gewinnen wir die abgetastete charakteristische Funktion $f(i, j)$ durch

$$u_1 = (i - 1)/SK - SK/2 \quad \text{mit } i = 1, 2, \dots, N \quad (4.30)$$

und

$$u_2 = (j - 1)/SK - SK/2 \quad \text{mit } j = 1, 2, \dots, N. \quad (4.31)$$

mit $N = SK^2$ gilt außerdem

$$u_{1\min} = u_{2\min} = -SK/2 \quad (4.32)$$

und
$$u_{1\max} = u_{2\max} = SK/2 \quad (4.33)$$

Durch dieses Verfahren erhalten wir die gleichen Abtastraten und Skalierungsfaktoren auch für die Wahrscheinlichkeitsdichte $p_Q(x_1, x_2)$.

Sie besteht dann aus N^2 Abtastwerten (N Werte in jeder Dimension) im Bereich von

$$x_{1\min} = x_{2\min} = -SK/2 \quad (4.34)$$

und
$$x_{1\max} = x_{2\max} = SK/2 \quad (4.35)$$

Um den benötigten Speicherplatz zu begrenzen, wurde $ISK = 3$ gewählt; daraus ergibt sich ein Skalierungsfaktor von $SK = 8$ und eine Anzahl von $N^2 = 64^2 = 4096$ komplexen Abtastwerten (siehe Bild 4.2.a und Bild 4.2.b). Der Koordinatenursprung liegt in diesem Fall in der Mitte der Graphik, d.h. bei $i = j = N/2 + 1$.

Im nächsten Schritt werden die Abtastwerte mit einer zweidimensionalen Fensterfunktion multipliziert, um die Fehler, die durch die Begrenzung von u auf $\pm SK/2$ entstehen, zu minimieren.

Es wurde das zweidimensionale Kaiser-Fenster benutzt; dieses ist durch den seinen Formfaktor α besonders anpassungsfähig. Es entsteht aus dem eindimensionalen Kaiser-Fenster

$$w(u) = \begin{cases} I_0(\alpha \sqrt{1 - (u/u_{\max})^2}) & \text{für } |u| \leq u_{\max} \\ 0 & \text{sonst} \end{cases} \quad (4.36)$$

durch zirkulare Rotation ([20], S. 119) mit

$$u = \sqrt{u_1^2 + u_2^2} \quad (4.37)$$

I_0 in (4.36) ist die modifizierte Besselfunktion erster Art nullter Ordnung:

$$I_0(x) = 1 + \sum_{n=1}^{\infty} \left[\frac{(x/2)^n}{k!} \right]^2 \quad (4.38)$$

Für den Formfaktor hat sich in vielen Fällen $\alpha = 4$ bewährt.

Die Bilder 4.3.a und 4.3.b. zeigen Real- und Imaginärteil nach der Fensterung mit $\alpha = 4$. Vor der Transformation ist nun noch eine Rotation des Koordinatenursprungs notwendig (die diskrete Fouriertransformation ist zyklisch). In Bild 4.2a und b. liegen die negativen Werte von u_1 und u_2 noch im Bereich von $i = 1, 2, \dots, N/2$, $j = 1, 2, \dots, N/2$ der Koordinatenursprung ($u_1=0$, $u_2=0$) bei ($i = N/2+1$, $j = N/2+1$) und die positiven Werte von u_1 und u_2 im Bereich von $i = N/2+2, N/2+3, \dots, N$, $j = N/2+2, N/2+3, \dots, N$

Bei der diskreten Fourier Transformation (DFT) liegt der Nullpunkt aber definitionsgemäß bei $i = j = 1$, die positiven Werte bei $i = 2, 3, \dots, N/2+1$, $j = 2, 3, \dots, N/2+1$ und die negativen bei $i = N/2+2, N/2+3, \dots, N$, $j = N/2+2, N/2+3, \dots, N$. Es ist also eine Rotation des Koordinatenursprungs von ($i = N/2+1$, $j = N/2+1$) nach ($i = 1$, $j = 1$)

notwendig, mit

$$\begin{array}{lll} f_R(i+N/2, j+N/2) = f(i, j) & \text{für } i=1, \dots, N/2 & j=1, \dots, N/2 \\ f_R(i-N/2, j+N/2) = f(i, j) & \text{für } i=N/2+1, \dots, N & j=1, \dots, N/2 \\ f_R(i+N/2, j-N/2) = f(i, j) & \text{für } i=1, \dots, N/2 & j=N/2+1, \dots, N \\ f_R(i-N/2, j-N/2) = f(i, j) & \text{für } i=N/2+1, \dots, N & j=N/2+1, \dots, N \end{array} \quad (4.39)$$

In Bild 4.4a und b sieht man die charakteristische Funktion nach der Rotation. Die diskrete Fourier Transformation wird nun mit Hilfe eines FFT-Programms von Norman M. Brenner [8] durchgeführt. Es wurde bereits 1967 veröffentlicht, wird aber nach eigenen Erfahrungen an Universalität und Geschwindigkeit auch von neuesten Programmentwicklungen kaum übertroffen; so sind k-dimensionale diskrete Fouriertransformationen mit beliebigem k und N_1 (auch Primzahlen) möglich. Natürlich wird in der Praxis die Anzahl der Dimensionen dadurch begrenzt, daß die Anzahl der Datenwerte $\prod_{i=1}^k N_i$ sehr schnell die Größe des Kernspeichers sprengt. Eine Auslagerung auf externe Massenspeicher ist in diesem Programm nicht vorgesehen.

Vor der graphischen Ausgabe der zweidimensionalen Wahrscheinlichkeitsdichte der nichtzentralen χ^2 -Verteilung ist es sinnvoll, eine weitere Rotation des Datenfeldes durchzuführen (das Bild 4.5 zeigt die zweidimensionale Wahrscheinlichkeitsdichte nach der Transformation), so daß der Koordinatenursprung von ($i = 1, j = 1$) wieder nach ($i = N/2 + 1, j = N/2 + 1$) rotiert wird. (Bild 4.6)

$$\begin{array}{lll}
 p(i+N/2, j+N/2) = p_R(i, j) & \text{für } i=1, \dots, N/2+1 & j=1, \dots, N/2+1 \\
 p(i-N/2, j+N/2) = p_R(i, j) & \text{für } i=N/2+1, \dots, N & j=1, \dots, N/2+1 \\
 p(i+N/2, j-N/2) = p_R(i, j) & \text{für } i=1, \dots, N/2+1 & j=N/2+1, \dots, N \\
 p(i-N/2, j-N/2) = p_R(i, j) & \text{für } i=N/2+1, \dots, N & j=N/2+1, \dots, N
 \end{array}
 \quad (4.40)$$

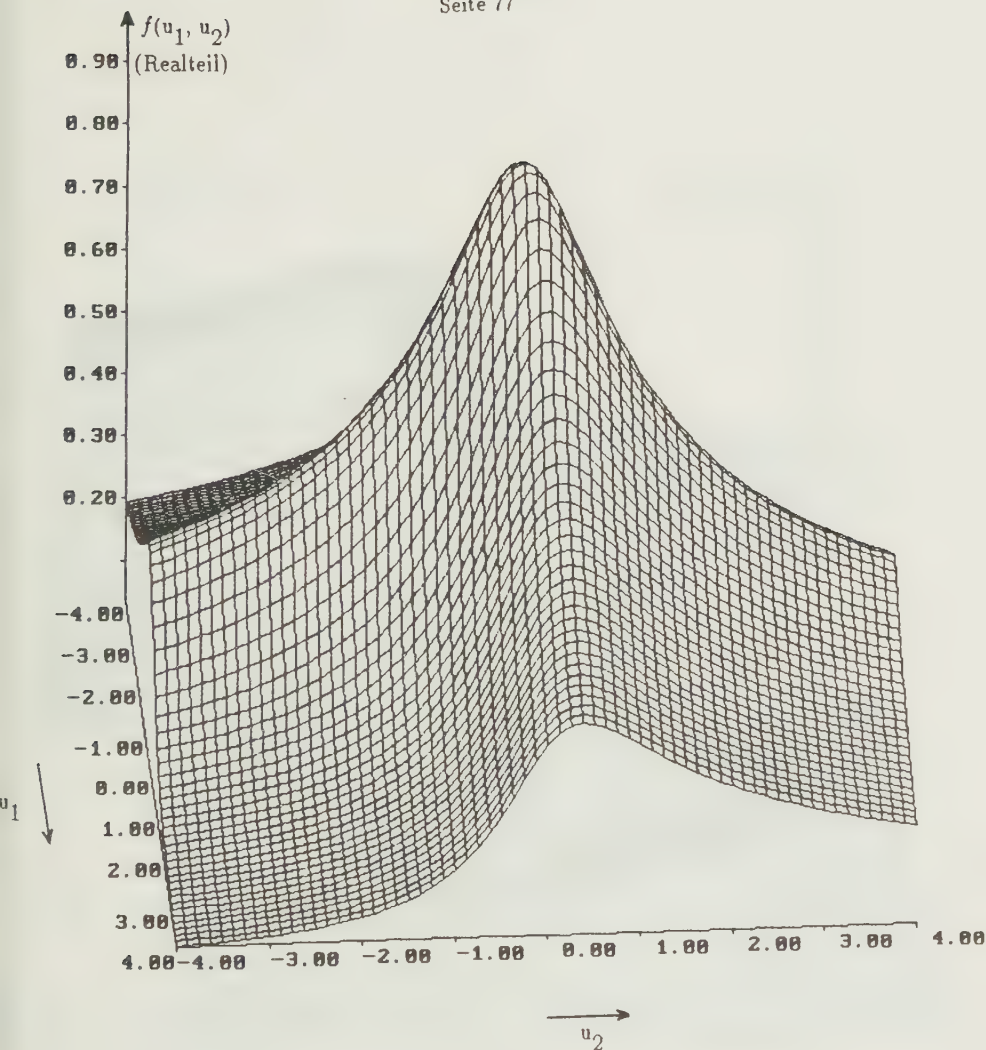


Bild 4.2.a. Charakteristische Funktion der zentralen χ^2 -Verteilung.
 Der Koordinatenursprung liegt bei $(N/2+1, N/2+1)$; die
 Korrelationskoeffizienten sind $R(0) = 0,1$ und $R(1) = 0$.

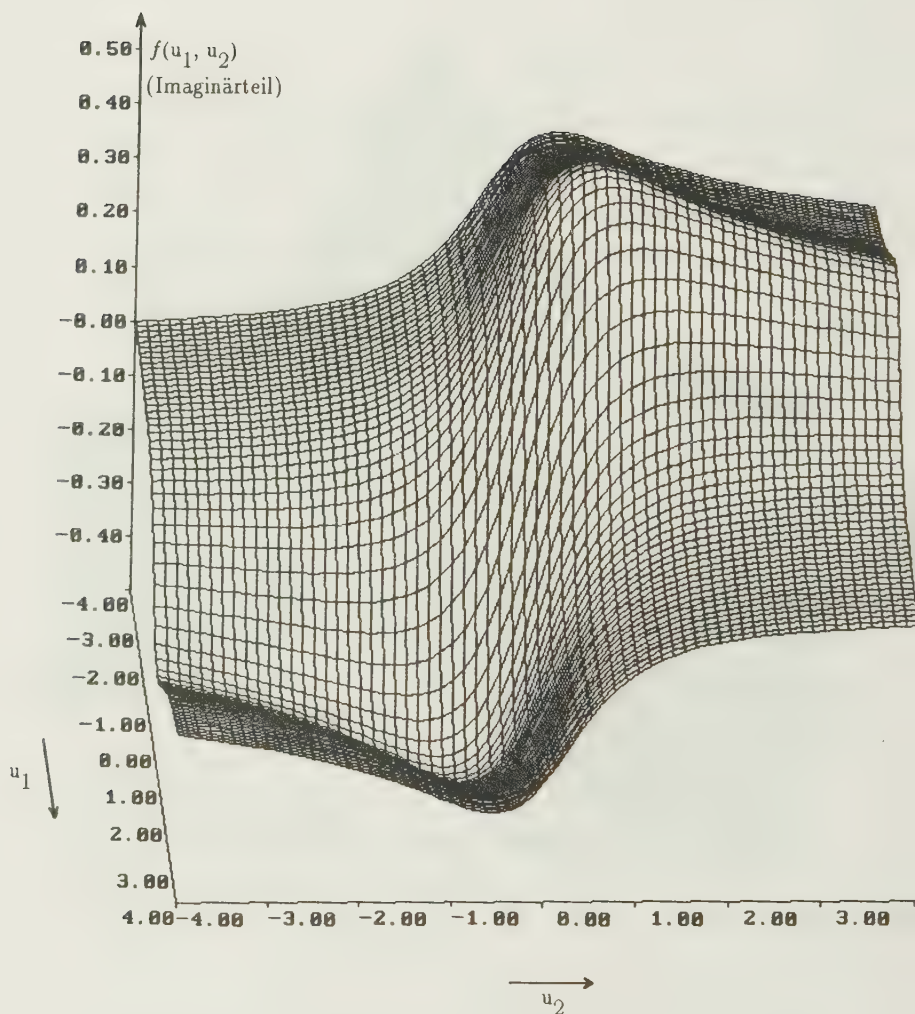


Bild 4.2.b. Charakteristische Funktion der zentralen χ^2 -Verteilung.

Der Koordinatenursprung liegt bei $(N/2+1, N/2+1)$; die

Korrelationskoeffizienten sind $R(0) = 0,1$ und $R(1) = 0$.

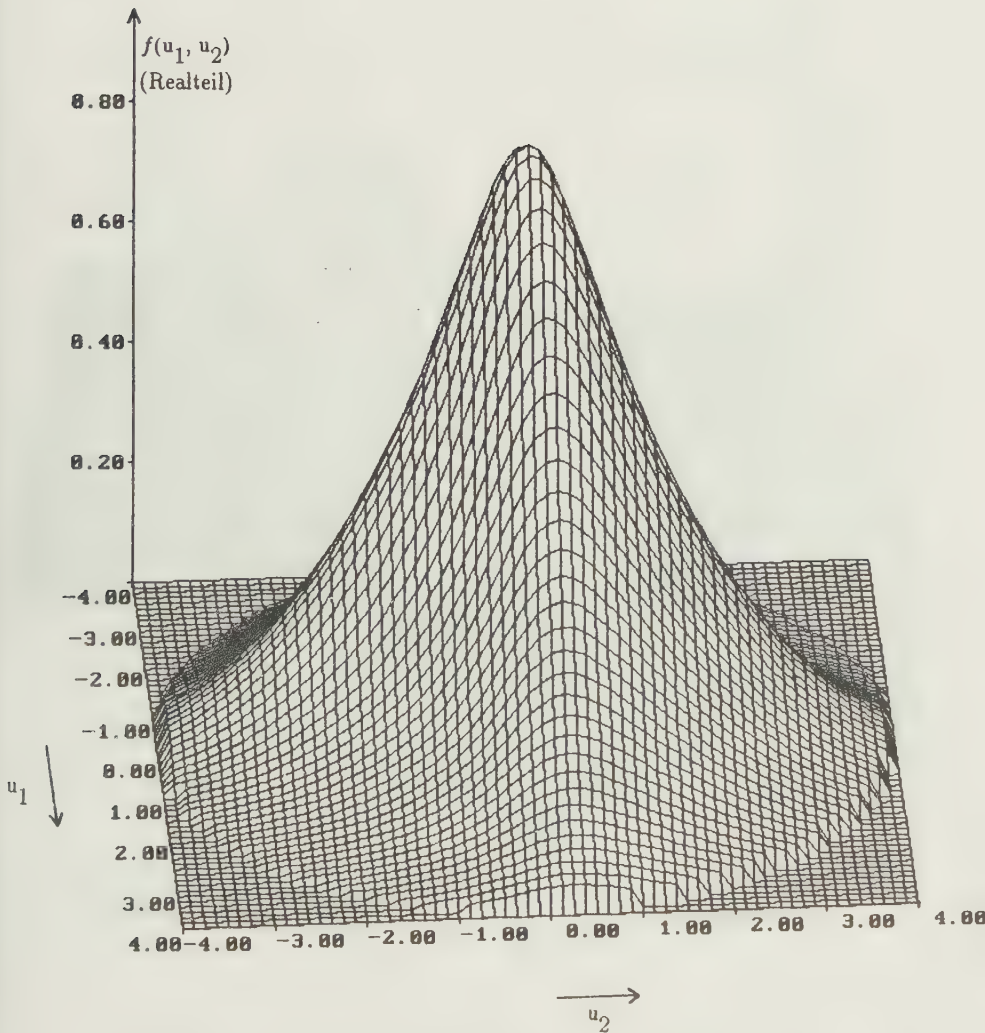


Bild 4.3.a. Charakteristische Funktion der zentralen χ^2 -Verteilung nach der Fensterung mit einem Kaiserfenster $\alpha = 4$.

Der Koordinatenursprung liegt bei $(N/2+1, N/2+1)$; die Korrelationskoeffizienten sind $R(0) = 0,1$ und $R(1) = 0$.

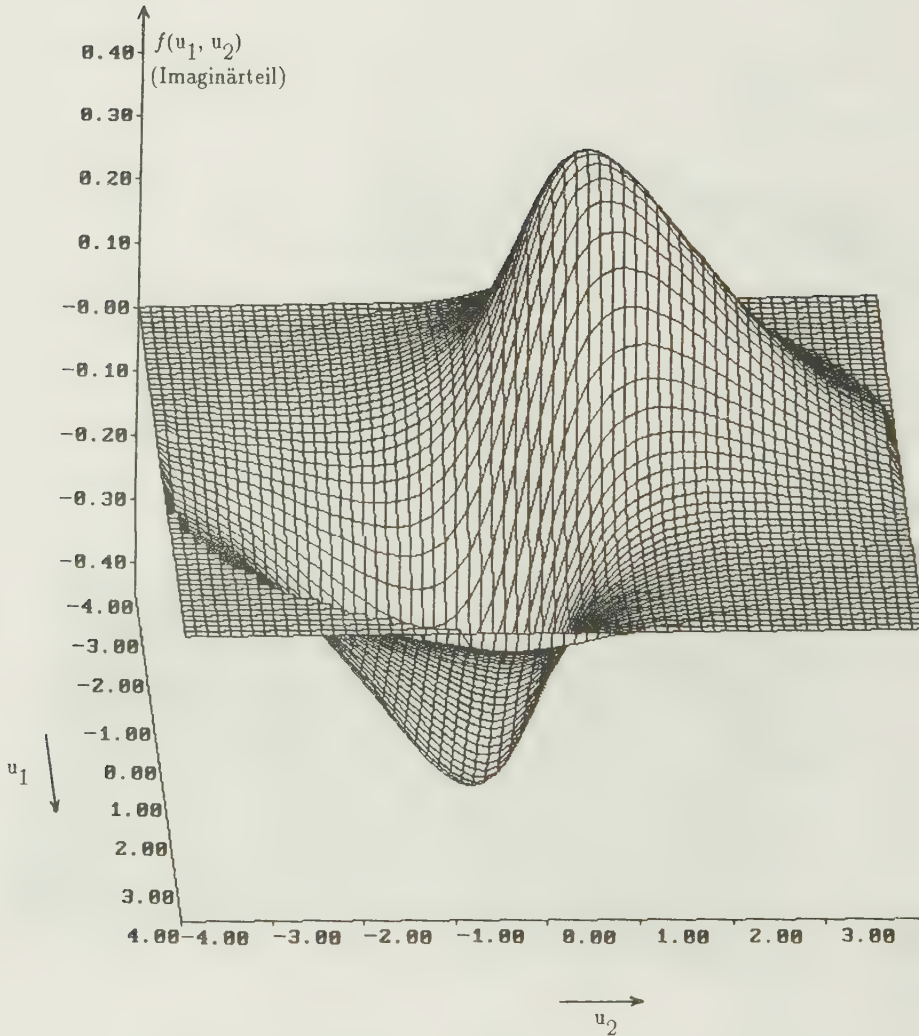


Bild 4.3.b. Charakteristische Funktion der zentralen χ^2 -Verteilung nach der Fensterung mit einem Kaiserfenster $\alpha = 4$.

Der Koordinatenursprung liegt bei $(N/2+1, N/2+1)$; die

Korrelationskoeffizienten sind $R(0) = 0.1$ und $R(1) = 0$

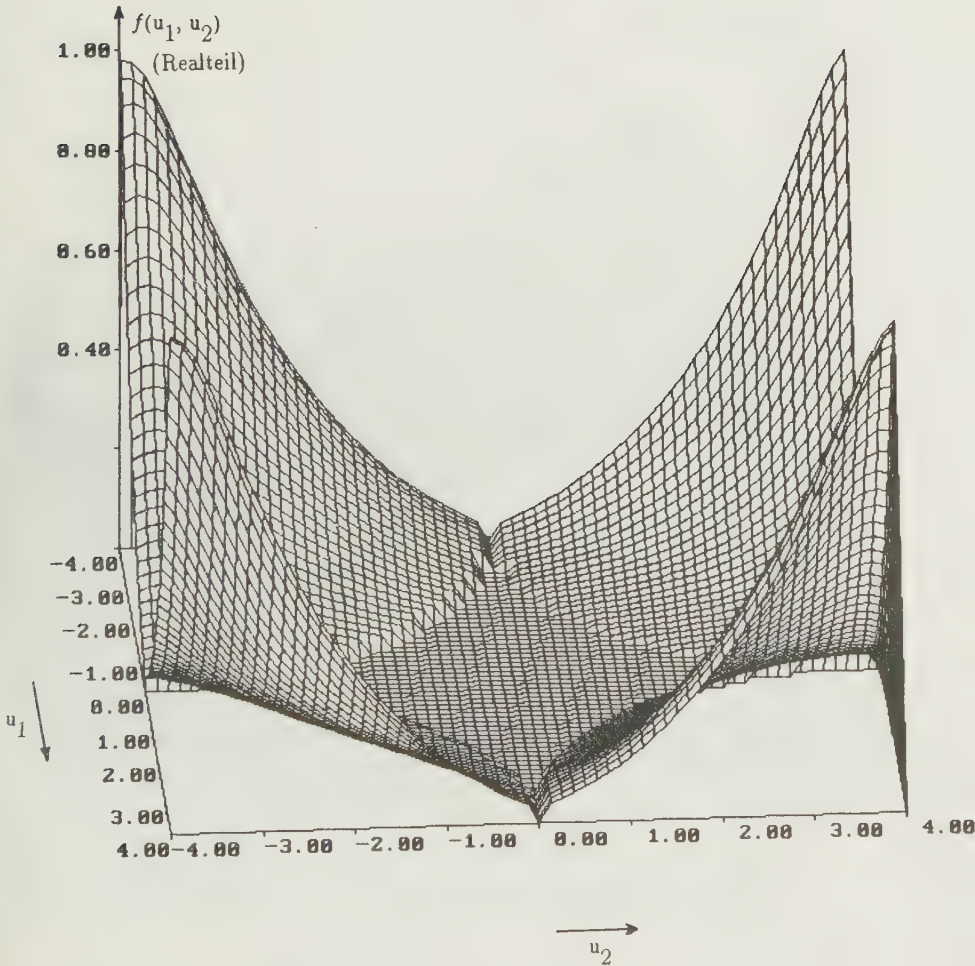


Bild 4.4.a. Charakteristische Funktion der zentralen χ^2 -Verteilung nach der Rotation.

Der Koordinatenursprung liegt bei (1, 1); die

Korrelationskoeffizienten sind $R(0) = 0,1$ und $R(1) = 0$.

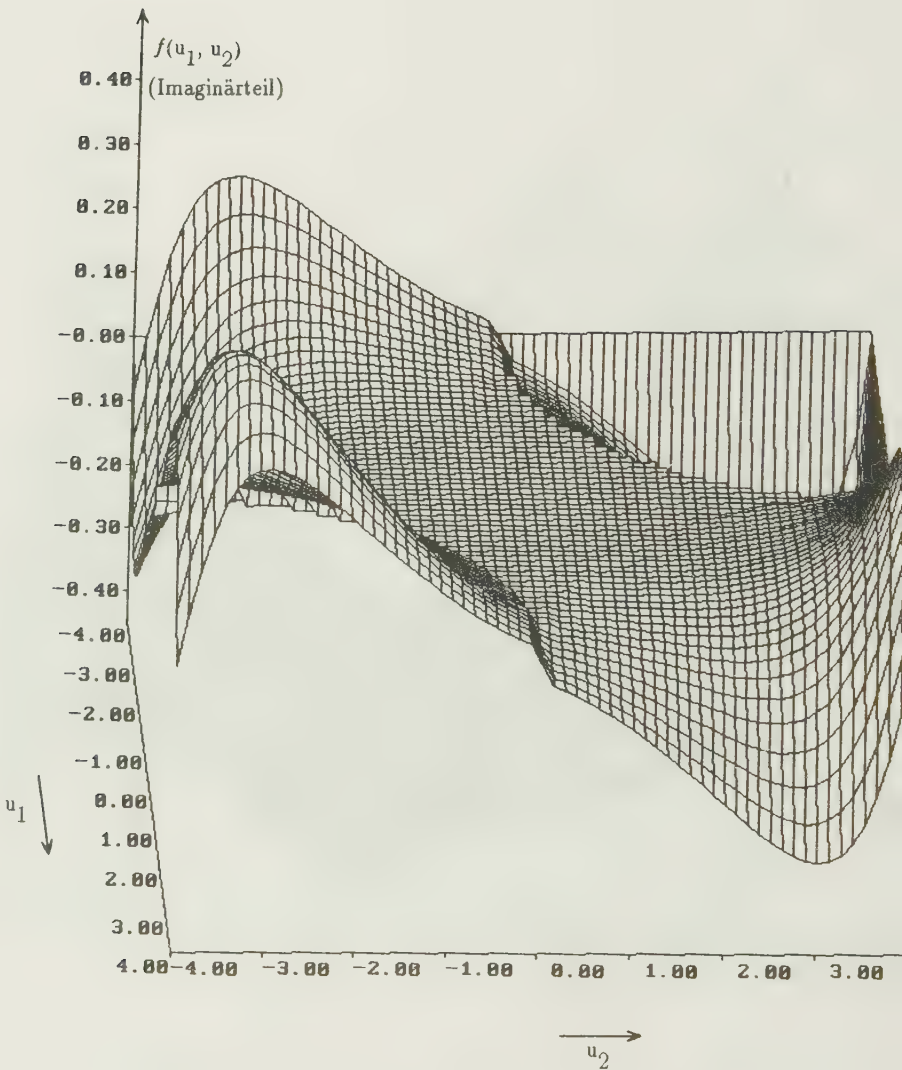


Bild 4.4.b. Charakteristische Funktion der zentralen χ^2 -Verteilung nach der Rotation.

Der Koordinatenursprung liegt bei (1, 1); die

Korrelationskoeffizienten sind

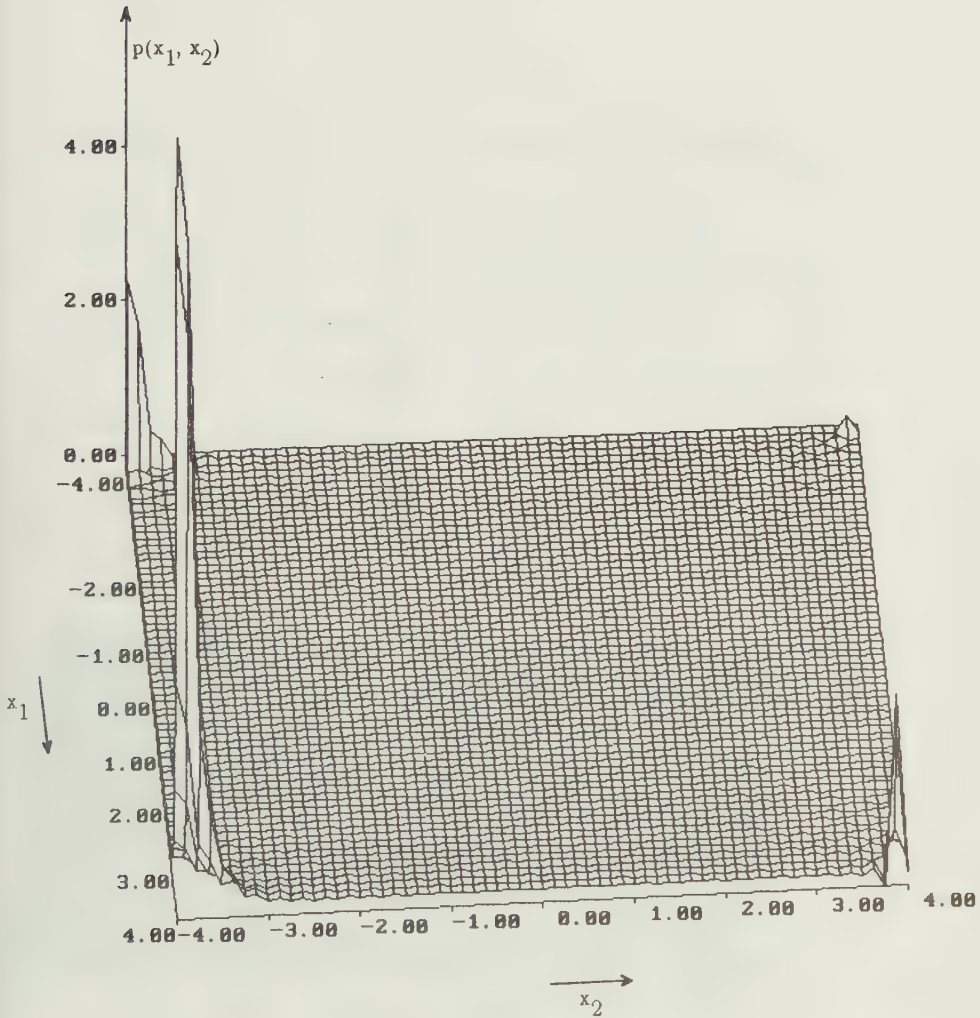


Bild 4.5. Wahrscheinlichkeitsdichte der zentralen χ^2 - Verteilung vor der Rotation.

Der Koordinatenursprung liegt bei (1, 1); die

Korrelationskoeffizienten sind $R(0) = 0,1$ und $R(1) = 0$.

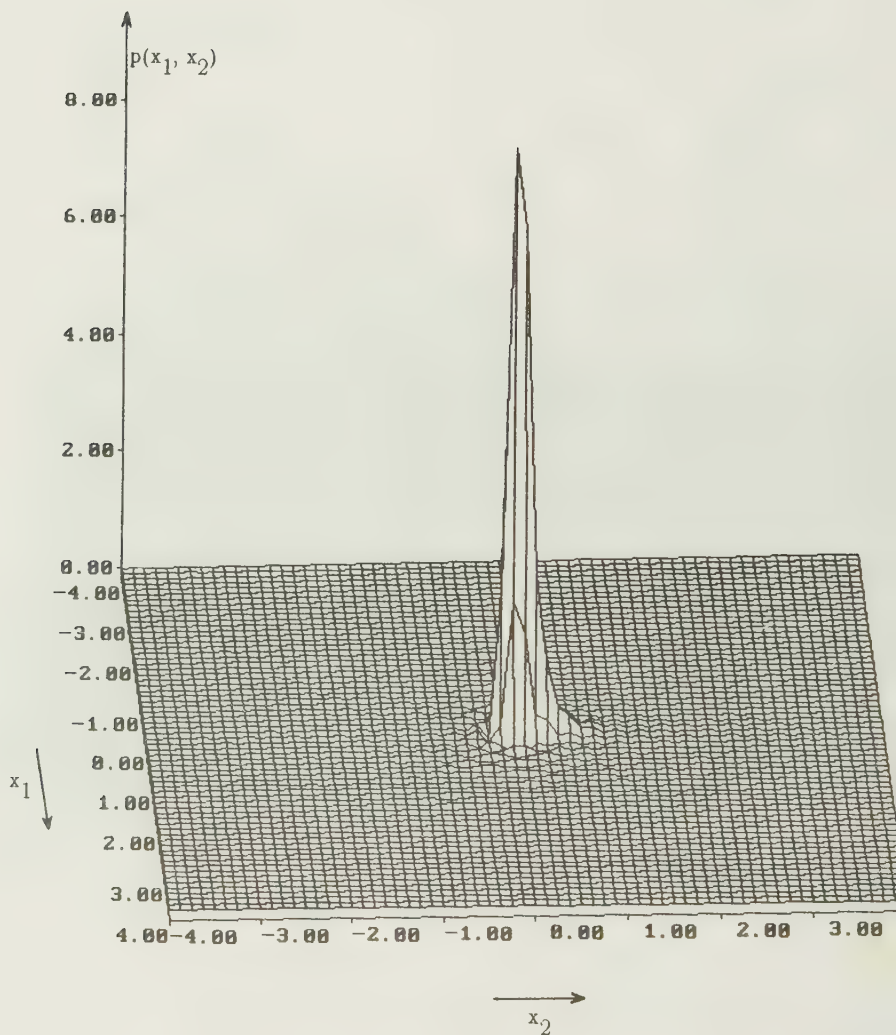


Bild 4.6. Wahrscheinlichkeitsdichte der zentralen χ^2 -Verteilung nach der Rotation.

Der Koordinatenursprung liegt bei $(N/2+1, N/2+1)$; die Korrelationskoeffizienten sind $R(0) = 0.1$ und $R(1) = 0$

Die Bilder 4.7 bis 4.12 zeigen zweidimensionale χ^2 -Verteilungen mit verschiedenen Signal-Rausch-Abständen und Korrelationen, außerdem den Fehler der von Jensen vorgeschlagenen Näherung (Satz 4.2.1).

Die zweidimensionale χ^2 -Verteilung mit der charakteristischen Funktion (Gleichung 4.23)

$$f_Q(u_1, u_2) = |\hat{I} - 2 j \hat{U}_D \hat{\Lambda}|^{-1/2} \cdot \exp \{ j \hat{S}_{ZF}^T (\hat{I} - 2 j \hat{U}_D \hat{\Lambda})^{-1} \hat{U}_D \hat{S}_{ZF} \}$$

ist vollständig determiniert durch die Varianz $\sigma^2 = \lambda_{11} = \lambda_{22}$, den Korrelationskoeffizienten $r(1) = \lambda_{12} / \lambda_{11} = \lambda_{21} / \lambda_{11}$ und den Signalvektor (s_1, s_2) .

Das Signal ist in allen Fällen $s_1 = s_2 = \sqrt{2}$. Der Korrelationskoeffizient (vor der Quadratur) ist in den Bildern 4.7 bis 4.9

$r(1) = 0$, das heißt, die Zufallsvariablen sind statistisch unabhängig. In den Bildern 4.10 bis 4.12 beträgt die Korrelation $r(1) = 0,8$.

Die Varianz, die der Rauschleistung vor der Quadratur entspricht, ist in den Bildern 4.7 und 4.10 $\sigma^2 = 1$, in den Bildern 4.8 und 4.11 $\sigma^2 = 0,1$ und in den Bildern 4.9 und 4.12 $\sigma^2 = 0,01$. Daraus ergibt sich ein Signal-Rausch-Abstand von 6, 16 und 26 dB.

Nach Satz 4.2.1 kann die nichtzentrale χ^2 -Verteilung angenähert werden durch eine Normalverteilung mit der Kovarianzmatrix (Gleichung 4.26)

$$k_{i,j} = \lambda_{i,j}^2 + 2 \cdot s_i \cdot s_j \cdot \lambda_{i,j}$$

und dem Mittelwert (Gleichung 4.25)

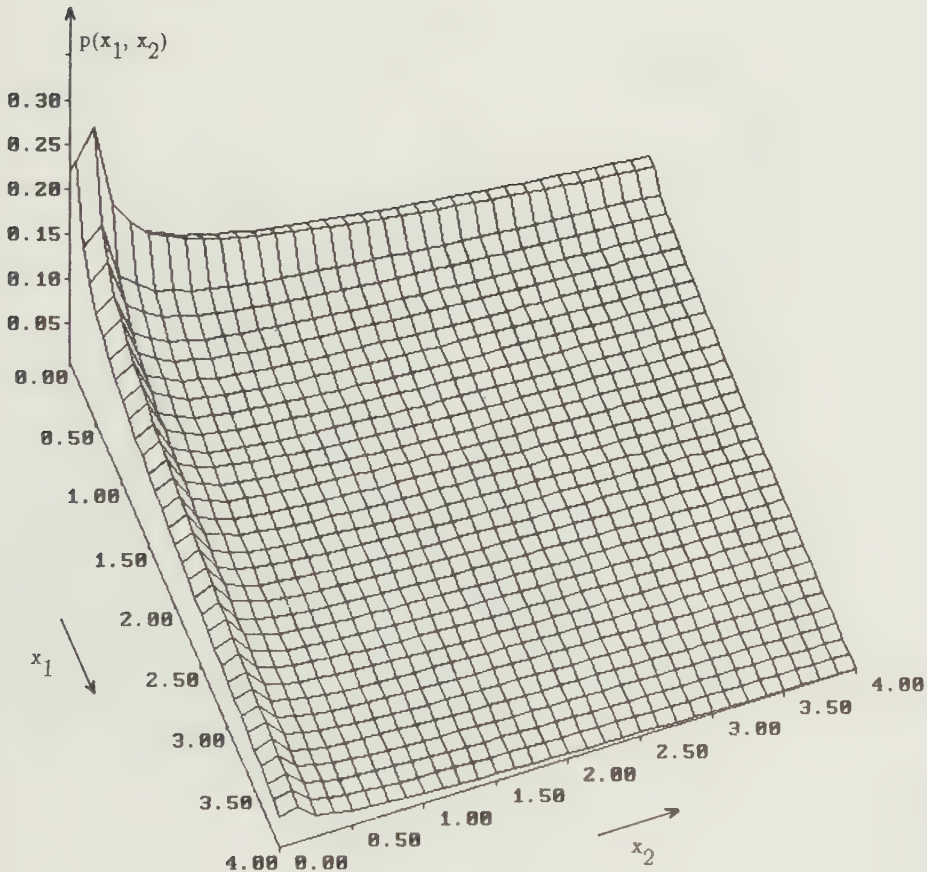
$$m_i = \lambda_{i,i} + s_i^2.$$

Diese Näherung gilt für große Signal-Rausch-Abstände (D. R. Jensen, [32]). Die Bilder 4.7.b, 4.8.b, 4.9.b, 4.10.b, 4.11.b und 4.12.b zeigen die Differenz zwischen der nichtzentralen χ^2 -Verteilung und der von Jensen vorgeschlagenen Näherung durch die Normalverteilung.

Man erkennt, daß der Fehler mit größer werdendem Signal-Rausch-Abstand schnell abnimmt. Bei einem Signal-Rausch-Abstand von 6 dB ist die Normalverteilung jedoch eine denkbar schlechte Näherung, wie man in den Bildern 4.7.a und b sowie 4.10.a und b sehen kann. Der Fehler ist fast so groß wie die Verteilung selbst.

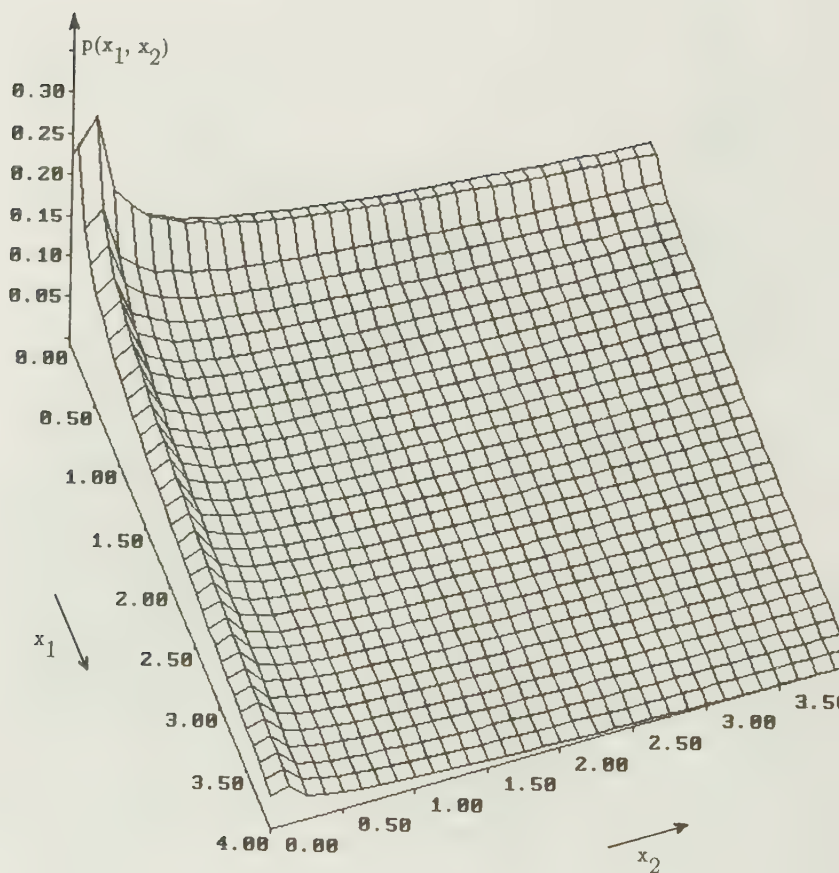
Für hohen Signal-Rausch-Abstand (16 und 26 dB) sieht die Verteilung einer Normalverteilung schon sehr ähnlich; der Fehler ist entsprechend gering (Bild 4.8.b, 4.9.b, 4.11.b und 4.12.b).

In den Bildern 4.11.a und 4.12.a erkennt man bereits die für korrelierte und normalverteilte Zufallsvariablen typische Konzentration der Verteilung auf die Diagonale.



4.7.a. Nichtzentrale χ^2 - Verteilung.

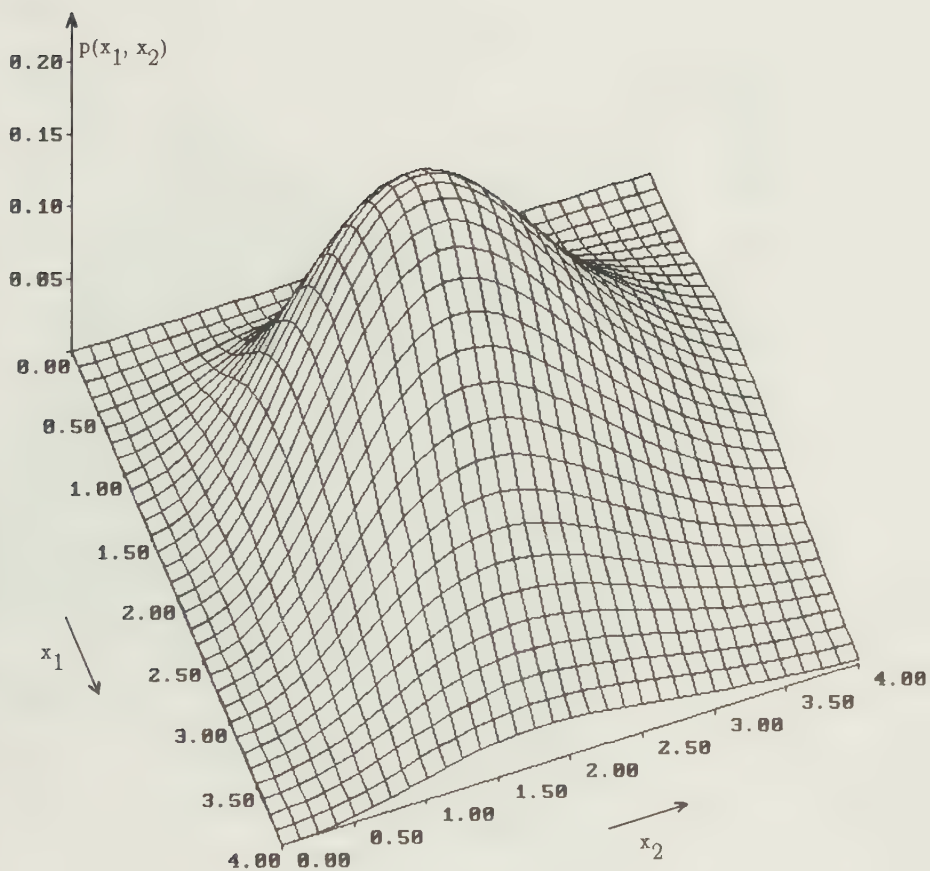
Der Signal-Rausch-Abstand beträgt 6 dB,
 der Korrelationskoeffizient ist 0,0.



4.7.b. Differenz zwischen der χ^2 -Verteilung und der entsprechenden Normalverteilung.

Der Signal-Rausch-Abstand beträgt 6 dB,

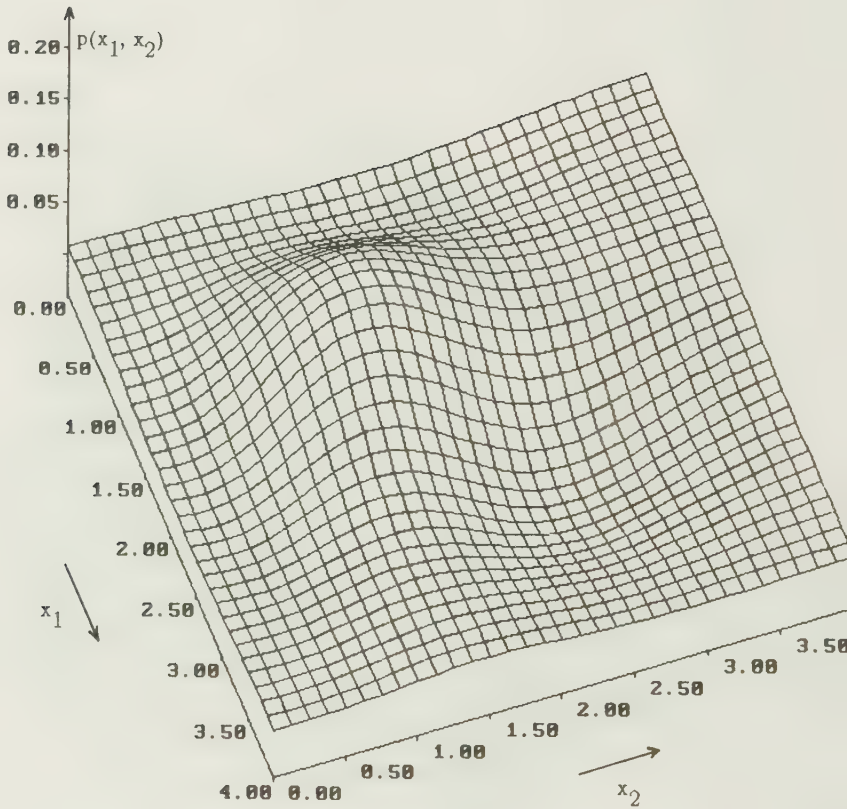
der Korrelationskoeffizient ist 0,0.



4.8.a. Nichtzentrale χ^2 - Verteilung.

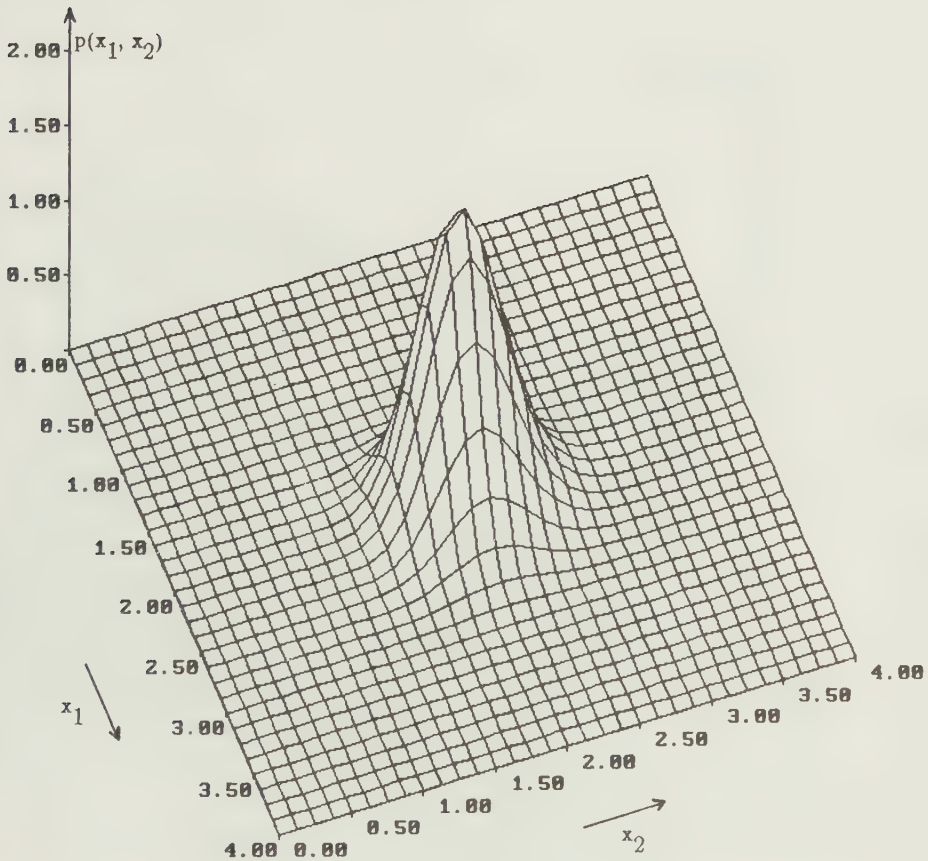
Der Signal-Rausch-Abstand beträgt 16 dB,

der Korrelationskoeffizient ist 0,0.



4.8.b. Differenz zwischen der χ^2 -Verteilung und der entsprechenden Normalverteilung.

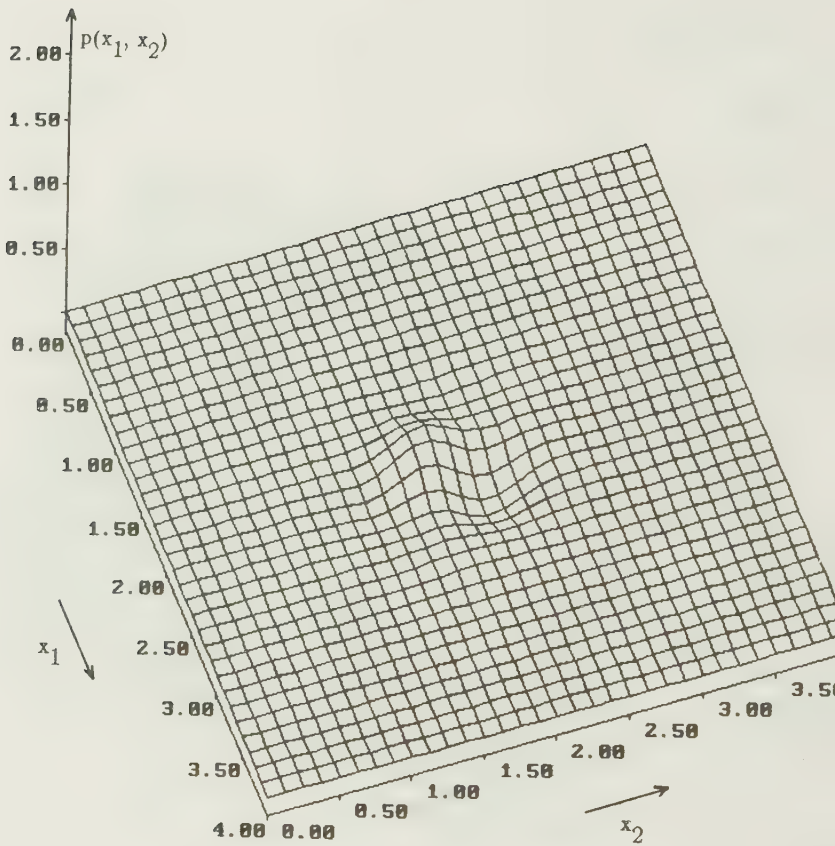
Der Signal-Rausch-Abstand beträgt 16 dB,
der Korrelationskoeffizient ist 0,0.



4.9.a. Nichtzentrale χ^2 - Verteilung.

Der Signal-Rausch-Abstand beträgt 26 dB,

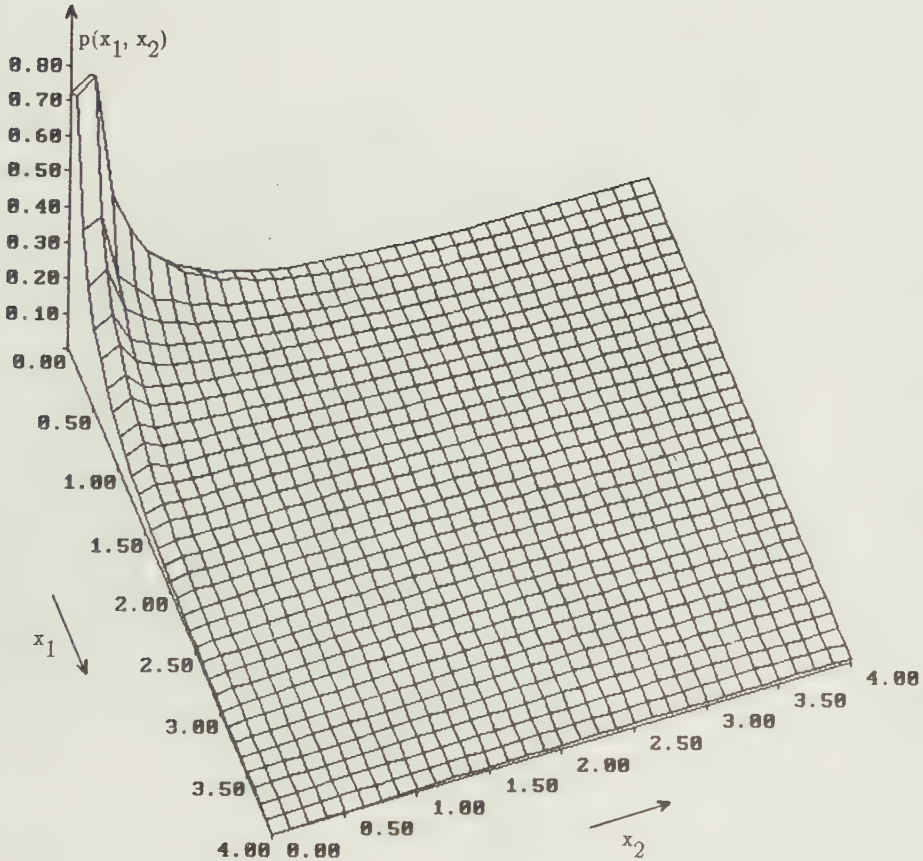
der Korrelationskoeffizient ist 0,0.



4.9.b. Differenz zwischen der χ^2 - Verteilung und der entsprechenden Normalverteilung.

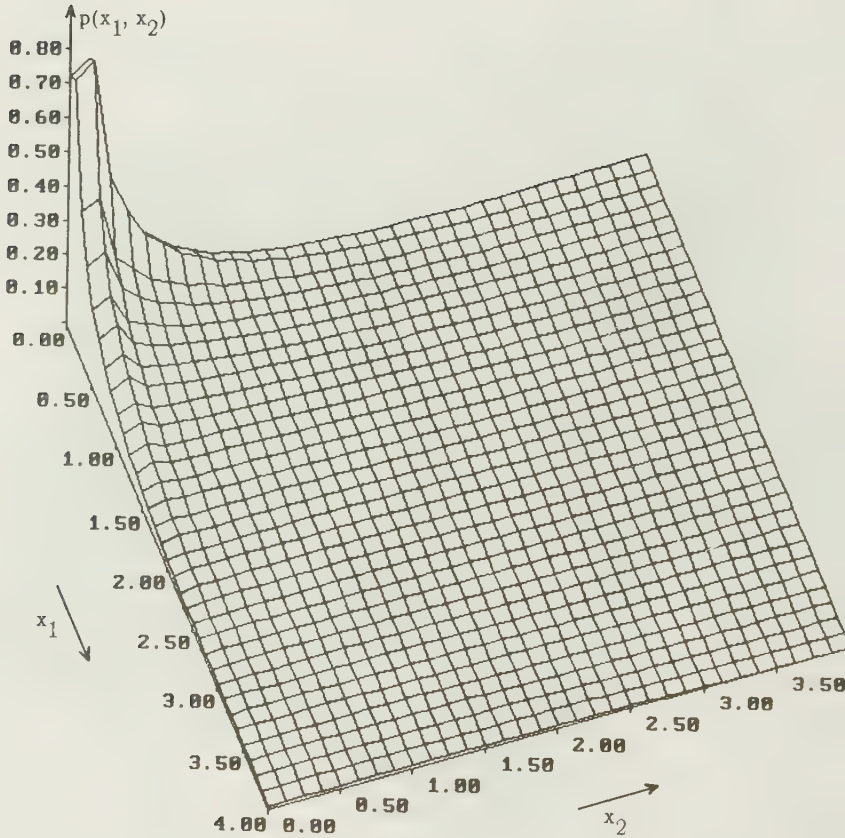
Der Signal-Rausch-Abstand beträgt 26 dB,

der Korrelationskoeffizient ist 0,0.



4.10.a. Nichtzentrale χ^2 - Verteilung.

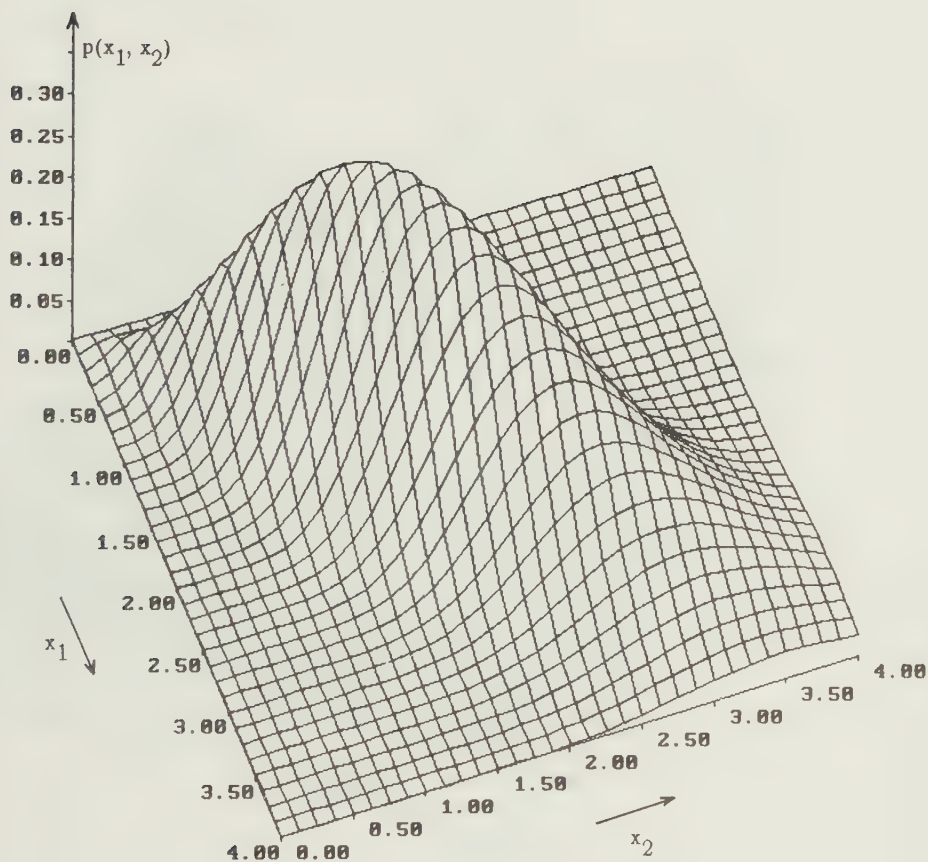
Der Signal-Rausch-Abstand beträgt 6 dB,
 der Korrelationskoeffizient ist 0,8.



4.10.b. Differenz zwischen der χ^2 - Verteilung und der
entsprechenden Normalverteilung.

Der Signal-Rausch-Abstand beträgt 6 dB,

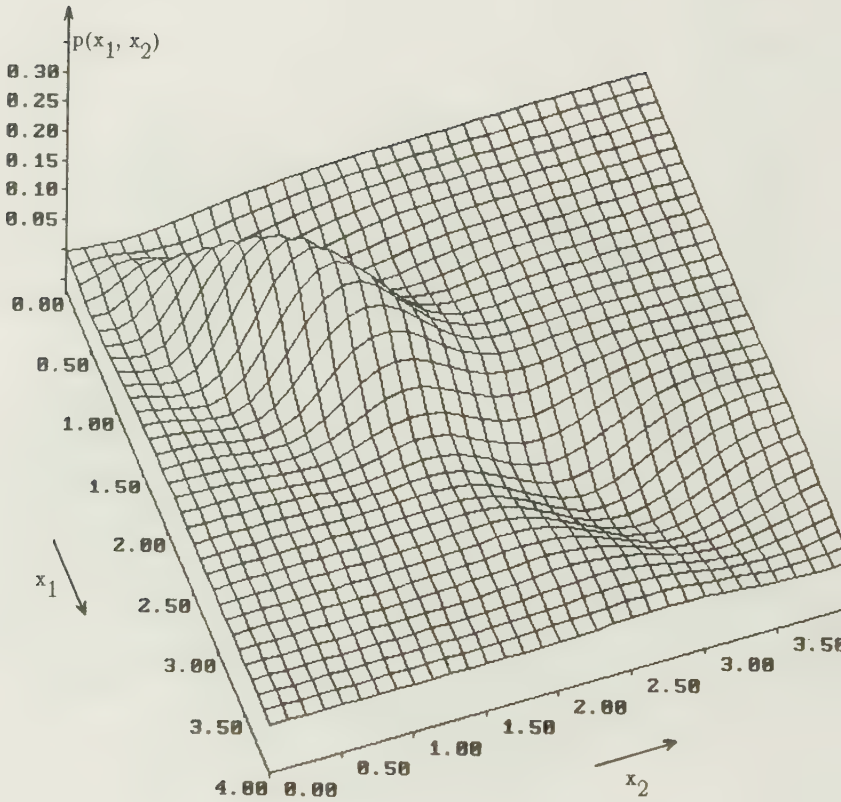
der Korrelationskoeffizient beträgt



4.11.a. Nichtzentrale χ^2 - Verteilung.

Der Signal-Rausch-Abstand beträgt 16 dB

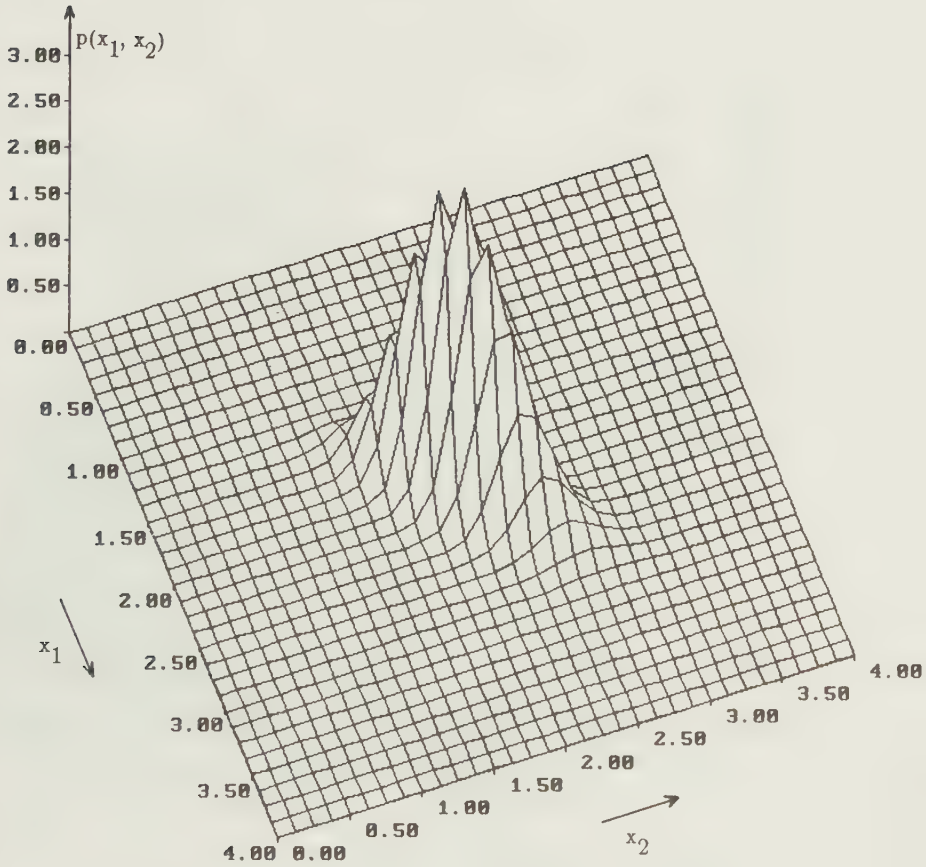
der Korrelationskoeffizient ist 0,8.



4.11.b. Differenz zwischen der χ^2 - Verteilung und der entsprechenden Normalverteilung.

Der Signal-Rausch-Abstand beträgt 16 dB

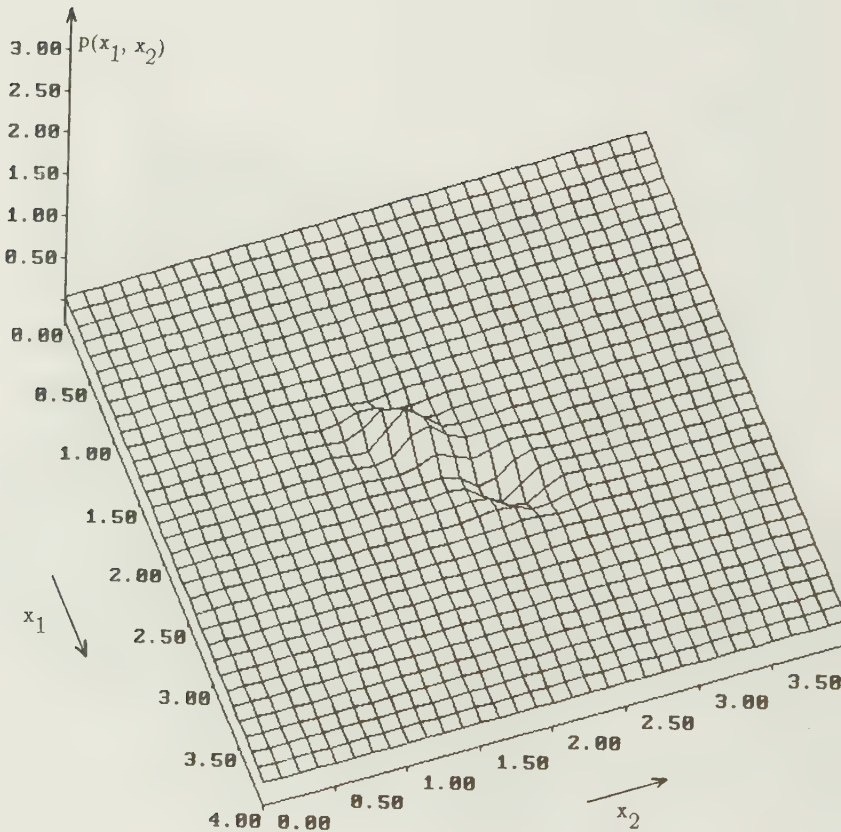
der Korrelationskoeffizient ist 0,8.



4.12.a. Nichtzentrale χ^2 - Verteilung.

Der Signal-Rausch-Abstand beträgt 26 dB,

der Korrelationskoeffizient ist 0,8.



4.12.b. Differenz zwischen der χ^2 - Verteilung und der entsprechenden Normalverteilung.

Der Signal-Rausch-Abstand beträgt 26 dB,

der Korrelationskoeffizient ist 0,8.

4.3 Wahrscheinlichkeitsdichte am Ausgang des NF-Filters

Eine geschlossene Lösung für die charakteristische Funktion des stochastischen Prozesses am Eingang des linearen NF-Filters liegt nun vor (Gleichung 4.23). Mit der in Kapitel 3.6 erarbeiteten Methode ist es daher möglich, eine geschlossene Lösung für die eindimensionale charakteristische Funktion nach der Filterung anzugeben, falls die Impulsantwort des linearen Filters die Länge k hat (Gleichung 3.59).

Fassen wir kurz zusammen:

Die gemeinsame Verteilungsfunktion von k aufeinanderfolgenden Abtastwerten der Summe von Signal und normalverteiltem weißem Rauschen nach der ZF-Filterung ist die k -dimensionale Normalverteilung mit der charakteristischen Funktion

$$f_{ZF}(\vec{u}) = \exp\{j \cdot \vec{u}^T \cdot \vec{s}_{ZF} - \frac{1}{2} \sigma^2 \cdot \vec{u} \cdot \vec{\Lambda} \cdot \vec{u}^T\}. \quad (4.41)$$

Nach der Quadratur haben wir einen k -dimensionalen nichtzentral χ^2 -verteilten Zufallsvektor $\vec{x}_Q$ für dessen charakteristische Funktion eine geschlossene Lösung bekannt ist, mit

$$f_Q(\vec{v}) = |\vec{I} - 2j \vec{v}_D^T \vec{\Lambda}|^{-1/2} \cdot \exp\{j \vec{s}_{ZF}^T (\vec{I} - 2j \vec{v}_D^T \vec{\Lambda})^{-1} \vec{v}_D \vec{s}_{ZF}\} \quad (4.42)$$

Mit Hilfe von Satz 3.6.1 erhalten wir nun eine Näherungslösung k -ter Ordnung für die eindimensionale charakteristische Funktion nach der NF-Filterung, indem wir die Impulsantwort im Punkt k abbrechen:

$$\vec{h} = (h(k-1), \dots, h(1), h(0)) \quad (4.43)$$

Ist $f_{\vec{X}}(\vec{v})$ die k -dimensionale charakteristische Funktion vor der linearen Filterung, dann ist die Näherung für die eindimensionale charakteristische Funktion nach der Filterung gegeben durch

$$f_1^k(u) = f_{\vec{X}_1}(\vec{v} = u \cdot \vec{h}) = f_{\vec{X}_1}(u \cdot h(k-1), \dots, u \cdot h(1), u \cdot h(0)) \quad (4.44)$$

Die Wahrscheinlichkeitsdichte ist die Fouriertransformierte der charakteristischen Funktion (Satz 3.3.10) mit

$$p_1^k(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-juy) \cdot f_1^k(u) du. \quad (4.45)$$

Eine exakte Lösung erhalten wir dann in Form eines Grenzwertes

$$p_1(y) = \lim_{k \rightarrow \infty} p_1^k(y) \quad (4.46)$$

Mit (4.42) und (4.44) ist nun die eindimensionale charakteristische

Funktion am Ausgang des Empfängers:

$$f_{NF,i}^k(u) = |\hat{I} - 2j u \hat{H}_D \hat{\Lambda}|^{-1/2} \cdot \exp \{j u \hat{S}_{ZF}^T (\hat{I} - 2j u \hat{H}_D \hat{\Lambda})^{-1} \hat{H}_D \hat{S}_{ZF}\} \quad (4.47)$$

Dabei ist $\hat{H}_D$ eine Diagonalmatrix mit Elementen, die den ersten k Werten der Impulsantwort entsprechen.

$$\hat{H}_D = \begin{bmatrix} h(k-1) & 0 & \cdots & 0 \\ 0 & h(k-2) & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & h(0) \end{bmatrix} \quad (4.48)$$

Während $\hat{\Lambda}$ die Kovarianzmatrix des Zufallsvektors $\hat{x}_{ZF}$ vor der Quadratur, mit

$$\lambda_{i,j} = R_N(\ell = i-j) \quad i, j = 1, 2, \dots, k \quad (4.49)$$

und

$$R_N(\ell) = \sigma^2 \cdot [g(\ell) * g(-\ell)] \quad (4.50)$$

ist. σ^2 ist hier die Rauschleistung am Eingang des ZF - Filters, und den Vektor $\hat{S}_{ZF} = (S_{ZF}(i-k+1), S_{ZF}(i-k+2), \dots, S_{ZF}(i-1), S_{ZF}(i))$ erhält man durch die Faltung des deterministischen Eingangssignals mit der Impulsantwort des ZF-Filters:

$$S_{ZF}(i) = g(i) * S_E(i), \quad i = 1, 2, \dots, k \quad (4.51)$$

5. Numerische Berechnung

5.1 Vorüberlegungen

Es soll nun ein effektives numerisches Verfahren zur Berechnung der Wahrscheinlichkeitsdichte am Ausgang des Empfängers entwickelt werden. Mit Gleichung (4.45) können wir für die Wahrscheinlichkeitsdichte schreiben:

$$p_1^k(y) = \mathcal{F} \{f_1^k(u)\}. \quad (5.1)$$

Dabei ist $\mathcal{F}$ der Operator der Fouriertransformation und $f_1^k(u)$ die charakteristische Funktion am Ausgang des Empfängers. Diese wird nun nach Gleichung (4.47) entwickelt:

$$f_1^k = d^{-\frac{1}{2}} \cdot \exp(j \cdot q) \quad (5.2)$$

die Terme d und q entsprechen

$$q = u \hat{S}_{ZF}^T \hat{C} \hat{H}_D \hat{S}_{ZF}, \quad (5.3)$$

$$\text{mit} \quad \hat{C} = \hat{\Lambda}^{-1} \quad (5.4)$$

$$\text{und} \quad d = \det | \hat{\Lambda} |. \quad (5.5)$$

$$\text{Außerdem ist} \quad \hat{\Lambda} = \hat{I} - 2 j u \hat{H}_D \hat{\Lambda} \quad (5.6)$$

$\hat{\Lambda}$ erhalten wir aus Gleichung (4.49) und (4.50)

5.2 Autokovarianzmatrix

Die Autokorrelationsfunktion $R(\ell)$ aus Gleichung (4.50) kann numerisch nur für Impulsantworten mit endlicher Länge berechnet werden. Für Filter mit unendlicher Impulsantwort erhält man die Näherungslösung durch Abbruch der Reihe bei N .

Eine schnelle Berechnung ist mittels FFT möglich, indem man die Impulsantwort in den Fourier-Bildbereich transformiert und das Quadrat des Betrages durch inverse Transformation wieder in den Originalbereich bringt. (E. O. Brigham [10], S.241) Man muß jedoch bedenken, daß die diskrete Fourier Transformation (DFT) zyklisch ist und $R(\ell)$ eine Länge von $2N$ hat; um Überlappungen zu vermeiden, sind an $g(i)$ deshalb mindestens N Nullen anzuhängen. Der benutzte FFT-Algorithmus ist die "Duhamel-Hollmann-Split-Radix-FFT" [60], die in Kapitel 5.7 beschrieben wird.

$$R(\ell) = \text{DFT}^{-1} \left(|G(\ell)|^2 \right) \quad \ell = 1, 2, \dots, 2N \quad (5.7)$$

$$|G(\ell)|^2 = \text{Re}[G(\ell)]^2 + \text{Im}[G(\ell)]^2 \quad \ell = 1, 2, \dots, 2N \quad (5.8)$$

$$G(n) = \text{DFT}(g(n)) \quad n = 1, 2, \dots, 2N \quad (5.9)$$

Der Rechenzeitaufwand des Unterprogramms "FAKF" ist proportional zu $2N \cdot \log(2N)$.

5.3 Die komplexe Matrix A

Das Unterprogramm "I_2juHdL" berechnet die komplexe Matrix $\hat{A}$ nach Gleichung (5.6). Mit Gleichung (4.49) sind ihre Elemente

$$a_{i,j} = 1 - 2 \cdot j \cdot u \cdot R(0) \cdot h(i) \quad \text{für } i=j \quad (5.10)$$

$$a_{i,j} = -2 \cdot j \cdot u \cdot R(i-j) \cdot h(i) \quad \text{für } i \neq j \quad (5.11)$$

wobei $R(\ell)$ in Kapitel 5.2 ermittelt wurde. Rechenzeit $\sim k^2$.

5.4 Die Berechnung der Kehrmatrix und der Determinante

Die Berechnung von Gleichung (5.4) und (5.5) wird mit Hilfe der Gaußelimination im Unterprogramm "CINVMA" durchgeführt.

$\hat{A}$ wird zunächst auf eine obere Dreiecksform gebracht; die Determinante d erhält man dann durch Multiplikation der Diagonalelemente (Zurmühl 1965 S. 118) und die Kehrmatrix $\hat{C}$ durch Vertauschung der Elemente. Besonders stabil wird diese Methode durch Pivotisierung, d.h. dadurch, daß man als Eliminationszeile diejenige mit dem betragsgrößten Element benutzt.

Ein sehr gut strukturiertes FORTRAN-77-Programm mit diesem Algorithmus wurde mir freundlicherweise von Joachim Dittloff zur Verfügung gestellt. [17]

Der Rechenaufwand ist proportional zu k^3 .

5.5. Der Zentralitätsparameter

entspricht in unserem Fall dem Signalvektor $\hat{S}_{ZF}$; sind alle Elemente Null, so erhalten wir die charakteristische Funktion mit Gleichung (5.2) aus der Determinante $d = |\hat{A}|$ zu

$$f_1^k = d^{-\frac{1}{2}}. \quad (5.12)$$

andernfalls benötigen wir noch den Wert der quadratischen Form q , den wir mit Gleichung (5.2) berechnen:

$$q = u \cdot \sum_{i=1}^k \sum_{j=1}^k s_i c_{i,j} h_{jj} s_j \quad (5.13)$$

Dabei sind $c_{i,j}$ die Elemente der Kehrmatrix $\hat{C}_{h,jj}$, die Elemente der Diagonalmatrix $\hat{H}_D$ und s_i Elemente des Vektors $\hat{S}_{ZF}$.

Unterprogramm "usCHds" benötigt eine Rechenzeit proportional zu k^2 .

5.6 Abtastung und Fensterung

Die charakteristische Funktion $f(u)$ kann als inverse Fouriertransformierte der Wahrscheinlichkeitsdichte $p(x)$ aufgefaßt werden. Umgekehrt erhält man bei gegebener charakteristischer Funktion die Wahrscheinlichkeitsdichte

durch die Fouriertransformation

$$p(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-juy) \cdot f(u) du \quad (5.14)$$

unter der Bedingung, daß die zugehörige Verteilung $F(x)$ kontinuierlich ist. ([42], S. 33) Diese Formel kann auf k -dimensionale charakteristische Funktionen erweitert werden, deren Wahrscheinlichkeitsdichte man dann als k -dimensionale Fouriertransformierte erhält.

Für unsere Anwendung reicht jedoch die eindimensionale Transformation aus, da eine geschlossene Lösung für die eindimensionale charakteristische Funktion nach der Filterung vorliegt. Dennoch dürfte es kaum möglich sein, für unseren Anwendungsfall eine geschlossene Lösung für das Fourierintegral, d.h. für die Wahrscheinlichkeitsdichte am Ausgang des Empfängers, zu finden, abgesehen von einigen Spezialfällen.

So erhält man eine χ^2 - Verteilung mit k Freiheitsgraden, wenn man das ZF-Filter wegläßt und das abschließende NF-Filter aus k identischen Koeffizienten mit $h(n) = 1$ besteht.

Eine numerische Näherungslösung erhalten wir, indem wir die charakteristische Funktion diskretisieren (abtasten) und die kontinuierliche Fouriertransformation durch die diskrete Fouriertransformation (DFT) ersetzen. Diese kann dann mit Hilfe eines Algorithmus zur schnellen Fouriertransformation (Fast Fourier Transform : FFT) durchgeführt werden. Trotz der Einfachheit dieser Lösung ist die FFT bisher fast nie zur Berechnung der Wahrscheinlichkeitsdichte eingesetzt worden.

Häufiger findet man dagegen Reihenentwicklungen der Wahrscheinlichkeitsdichte mit Hilfe der Momente oder der Kumulanten (auch Semiinvarianten genannt). Das Moment n -ter Ordnung erhält man durch n -faches Differenzieren der charakteristischen Funktion an der Stelle $t = 0$; zur Berechnung der Kumulanten wird der Logarithmus der charakteristischen Funktion mehrfach differenziert.

Die Momentenmethode hat aber zwei Nachteile:

1. Bei mancher Verteilung existieren die Momente nur bis zu einer gewissen Ordnung; bei der Cauchy-Verteilung beispielsweise existiert kein einziges.
2. Im Gegensatz zur charakteristischen Funktion bestimmen die Momente die Dichte nicht immer eindeutig (Aussagen darüber, welchen Bedingungen die Momente genügen müssen, um eindeutig zu sein, findet man bei Kendall und Stuart 1977, S. 112, 113), d.h. bei einem gegebenen Satz von Momenten sind eventuell mehrere Lösungen für die Dichte möglich.

Vorteil der Momente ist, daß es mit ihrer Hilfe oft möglich ist, das asymptotische Verhalten der Dichte $p(x)$ für $x \rightarrow \infty$ zu bestimmen. Hier liegen eindeutig die Nachteile der Fourier Transformation mittels FFT.

Um eine Transformation auf dem Digitalrechner durchführen zu können, muß die zu transformierende Funktion sowohl im Original- als auch im Bildbereich diskret sein, und sie darf nur aus einer endlichen Anzahl von Datenwerten bestehen. Die diskrete Fouriertransformation erfüllt diese

Forderungen. Sie ist definiert als (Brigham [10], S. 123)

$$G\left(\frac{n}{N T}\right) = \sum_{k=0}^{N-1} g(k T) \exp(-2 j \pi n k / N) \quad (5.15)$$

Aus der Bedingung der Diskretheit der Funktion $g(k T)$ und ihrer Transformierten $G\left(\frac{n}{N T}\right)$ folgt, daß diese periodisch sein müssen.

$$G\left(\frac{n}{N T}\right) = G\left(\frac{r N + n}{N T}\right) \quad r = \dots -2, -1, 0, 1, 2, \dots \quad (5.16)$$

und

$$g(k T) = g[(r N + k) T] \quad r = \dots -2, -1, 0, 1, 2, \dots \quad (5.17)$$

Sind die Funktionen weder periodisch noch begrenzt in n (wie in unserem Fall), so kann die DFT dennoch als Näherung für die kontinuierliche Transformation benutzt werden, wobei Fehler aber unvermeidlich sind.

Eine weitere Konsequenz aus der Periodizität und der Definition der DFT ist die Lage der Funktionswerte für negative n bzw. k . Es gilt

$$G\left(-\frac{n}{N T}\right) = G\left(\frac{N - n}{N T}\right) \quad n = 1, 2, \dots N/2 \quad (5.18)$$

$$g(-k T) = g[(N - k) T] \quad n = 1, 2, \dots N/2 \quad (5.19)$$

Das Intervall $[(N/2), (N-1) T]$ enthält also bei der DFT die Funktionswerte für negative Argumente. Diese stellen für reelle Zeitsignale die negativen Frequenzanteile dar und werden meist nicht ausgewertet, da sie für reelle

Signale redundant sind (konjugierte Symmetrie). In unserem Fall ist die Funktion im Fourier Originalbereich (die charakteristische Funktion) komplex und im Bildbereich (Wahrscheinlichkeitsdichte) unsymmetrisch. Die diskreten Werte der Dichte werden deshalb von der Ausgabe mit Hilfe von Gleichung (5.18) umsortiert. Dies entspricht einer Rotation des Koordinatenursprungs.

Die Beschränkung auf N Datenwerte und die Diskretisierung mit einer Abtastrate von T führen zu den folgenden Fehlern, die sich durch Vergrößerung von N und Verkleinerung von T verringern lassen.

1. Die Beschränkung der Datenmenge auf N kann als Multiplikation mit einem rechteckigen Fenster aufgefaßt werden. Im Bildbereich führt dies zu einer Faltung mit der Transformierten des Fensters.

Die Größe des dadurch entstehenden Fehlers hängt von den Eigenschaften der Wahrscheinlichkeitsdichte ab. Besonders groß ist der Fehler bei Wahrscheinlichkeitsdichten, die einen Sprung enthalten (z.B. Exponential- oder χ^2 - Verteilungen); hier kommt es zu "Überschwingern", die als Gibbsches Phänomen bekannt sind.

Generell gilt: Je glatter die Wahrscheinlichkeitsdichte, desto geringer der Fehler, der durch Fensterung verursacht wird. Dichten ohne Sprünge oder schnelle Änderungen der Steigung haben charakteristische Funktionen, die für großes u schnell ausschlagen, wodurch der Fehler durch Fensterung gering bleibt.

Durch Wahl einer geeigneten Fensterfunktion (Rechteck ist meist ungünstig) kann man die Genauigkeit erhöhen. Das Kaiser-Fenster, das bereits bei

der zweidimensionalen Transformation benutzt wurde, hat sich für viele Anwendungen sehr bewährt:

$$w(k) = \begin{cases} I_0(\alpha \sqrt{1 - (k/N)^2}) & \text{für } |k| \leq N \\ 0 & \text{sonst} \end{cases} \quad (5.20)$$

I_0 ist auch hier wieder die modifizierte Besselfunktion erster Art nullter Ordnung (siehe Gleichung 4.38) und α ein Formfaktor, mit dem die Eigenschaften des Fensters beeinflusst werden können. Für $\alpha = 0$ geht das Kaiserfenster in ein Rechteckfenster über.

2. Die Diskretisierung (Abtastung) der charakteristischen Funktion mit der Abtastrate T führt zu einer Periodisierung der Dichte $p(x)$ mit der Periode T . Ist diese nicht durch das Intervall $(-T/2, T/2)$ beschränkt, kommt es zu Überlappungsfehlern (Alaising). Bei Wahrscheinlichkeitsdichten, die für $|x| \rightarrow \infty$ monoton fallen, hat der so entstehende Fehler bei $p(T/2)$ bzw. $p(-T/2)$ sein Maximum. Ist $p'(x)$ die durch die DFT gewonnene Näherung für die Wahrscheinlichkeitsdichte, so gilt für $x = T/2$:

$$p'(T/2) \approx p(T/2) + p(-T/2), \quad (5.21)$$

wenn man die Überlappung durch weitere Perioden vernachlässigt.

Für symmetrische Dichten gilt daher (Bild 5.1b)

$$p'(T/2) \approx 2 \cdot p(T/2) \quad (5.22)$$

Dieser Aliasingfehler ist für Anwendungen in der Stochastik der Hauptnachteil der DFT. Für viele Anwendungen ist das Verhalten der Verteilung von $p(x)$ für $|x| \rightarrow \infty$ von großer Bedeutung, gerade für große x ist der Fehler aber am größten. Bei symmetrischen Verteilungen ist der Wert der durch DFT ermittelten Dichte $p'(x)$ an den Intervallgrenzen $\pm T/2$ sogar mehr als doppelt so groß wie der wahre Wert. Bei der Auswertung der DFT-Lösung darf deshalb nur ein Bruchteil des Intervalls $[-T/2, T/2]$ ausgewertet werden, z.B. $[-T/2k, T/2k]$.

Durch dichtere Abtastung der charakteristischen Funktion, d.h. durch Verringerung der Abtastrate T , vergrößert sich das Intervall; damit ist es möglich, den absoluten Fehler beliebig zu verringern, da bei jeder Verteilung $p(x = T/2)$ durch Vergrößerung von x beliebig klein gemacht werden kann. Die Wahl des Abstandes $\Delta u = T$ zwischen den Abtastwerten der charakteristischen Funktion $f(u)$, d.h. die Wahl der Abtastrate T , hängt also von der gewünschten Genauigkeit ab und ist daher variabel gehalten worden. Analog zur zweidimensionalen Transformation wählen wir eine Skalierungszahl $ISK \in \mathbb{N}$ und mit dem Faktor

$$SK = 2^{ISK} \quad (5.23)$$

eine Abtastrate von

$$\Delta u = T = \frac{1}{SK} \quad (5.24)$$

Die Anzahl N der Datenwerte ist nun

$$N = SK^2 \quad (5.25)$$

und das Datenfenster ist das Intervall

$$[-SK/2, SK/2] \quad (5.26)$$

Der Abstand Δx der Datenwerte der Wahrscheinlichkeitsdichte ist bei dieser Wahl

$$\Delta x = \frac{1}{N \cdot T} = \frac{SK}{SK^2} = \frac{1}{SK} \quad (5.27)$$

Die Abtastrate ist also für $p(x)$ und $f(u)$ identisch, das gleiche gilt für die Fensterlänge. Mit der Definition der DFT (5.15) erhalten wir

$$p\left(\frac{n}{SK}\right) = \sum_{k=0}^{SK^2-1} f_w\left(\frac{k}{SK}\right) \exp(-2 j \pi n k / SK^2). \quad (5.28)$$

Dabei ist $f_w\left(\frac{k}{SK}\right)$ das Produkt aus der charakteristischen Funktion und der Fensterfunktion $w(k)$ nach Gleichung (5.20).

Da die charakteristische Funktion immer konjugiert symmetrisch ist (als inverse Fouriertransformation einer reellen Wahrscheinlichkeitsdichte) mit $f(-u) = f^*(u)$, genügt es, den positiven Ast der charakteristischen Funktion zu bestimmen, wodurch sich die Rechenzeit halbieren läßt.

Die Rechenzeit des Unterprogramms "PD1D" ist proportional zu N .

5.7 Duhamel-Hollmann-Split-Radix FFT

Mit der Vergrößerung der Anzahl N der Datenwerte nehmen die Rechenzeiten der diskreten Fouriertransformation nach Gleichung (5.15) überproportional zu. Eine direkte Berechnung der DFT erfordert für N Datenwerte N^2 Multiplikationen und Additionen.

Der 1965 von J.W. Cooley und J.W. Tukey [11] vorgestellte Algorithmus zur schnellen Fouriertransformation (FFT) benötigt dagegen weniger als $2 \cdot N \cdot \log_2(N)$ Multiplikationen und Additionen. Bei dieser Methode handelt es sich um einen Algorithmus zur Basis 2. Dieser kann immer dann eingesetzt werden, wenn die Anzahl N eine Potenz von 2 ist mit $N = 2^M$ $M \in \mathbb{N}$. Inzwischen sind aber Algorithmen bekannt, die anwendbar sind, wenn N aus beliebigen Faktoren zusammengesetzt ist.

Eine weitere Verringerung der Rechenzeit läßt sich durch eine Vergrößerung der Basis erzielen. Nachteil dieser Basis-4 oder Basis-8 Algorithmen [3] ist die Einschränkung der Anzahl der Datenwerte auf $N=4^M$ und $N = 8^M$. Die Anzahl der benötigten Multiplikationen sinkt aber auf etwas mehr als $N \cdot \log_2(N)$. Noch weniger Multiplikationen als der Basis-8 Algorithmus benötigt ein Verfahren, das 1984 von P. Duhamel und H. Hollmann [21] veröffentlicht wurde. Damit wird es zum ersten Mal möglich, $N = 2^M$ Werte mit weniger als $N \cdot \log_2(N)$ Multiplikationen zu transformieren.

Die "Duhamel-Hollmann-Split-Radix FFT" (SRFFT) entsteht durch Anwendung einer Basis-2 FFT auf Terme mit geradem Index und Anwendung von zwei Basis-4 FFTs auf Terme mit ungeradem Index.

Durch Einführung des Drehfaktors $W = e^{-j2\pi/N}$ können wir für die Definition der DFT (5.15) schreiben

$$G_n = \sum_{k=0}^{N-1} g_k W^{nk} \quad (5.29)$$

Grundidee der SRFFT ist die Aufteilung der DFT in

$$G_{2n} = \sum_{k=0}^{N/2-1} [g_k + g_{k+N/2}] W^{2nk} \quad (5.30)$$

für Terme mit geradem Index und

$$G_{4n+1} = \sum_{k=0}^{N/4-1} ([g_k - g_{k+N/2}] + j[g_{k+N/4} - g_{k+3N/4}]) W^k W^{4nk} \quad (5.31)$$

sowie

$$G_{4n+3} = \sum_{k=0}^{N/4-1} ([g_k - g_{k+N/2}] + j[g_{k+N/4} - g_{k+3N/4}]) W^{3k} W^{4nk} \quad (5.32)$$

für Terme mit ungeradem Index. Man erhält also eine DFT der Länge $N/2$ und zwei DFTs der Länge $N/4$ mit den zusätzlichen Drehfaktoren W^k (in Gleichung 5.31) und W^{3k} (in Gleichung 5.32). Setzt man dieses Verfahren rekursiv fort, indem man die drei DFTs mit derselben Methode weiter zerlegt, so erhält man den SRFFT-Algorithmus.

Weitere Redundanzen können beseitigt werden, wenn g_k im Original- oder G_n im Bildbereich reell ist. Bekanntlich kann man eine reelle Sequenz der Länge N durch eine komplexe DFT der Länge $N/2$ ersetzen. Im Bildbereich

müssen die Daten dann noch auf komplizierte Weise umsortiert und mit Drehfaktoren multipliziert werden (Brigham [10]S. 205), wodurch ein Teil des Rechenzeitgewinns wieder eingebüßt wird. Direkte Implementationen der FFT für reelle Daten sind schneller und einfacher zu benutzen. Das erste reelle FFT-Programm wurde bereits 1968 von Bergland veröffentlicht [4]; inzwischen wurde auch ein FORTRAN-77 Programm für die reelle Split-Radix-FFT entwickelt. Die Anzahl der benötigten Multiplikationen wird dadurch noch einmal um den Faktor 2 auf unter $N/2 \log_2(N)$ reduziert (H.V. Sorensen, M.T. Heideman, C.S. Burrus, [61]). Das erwähnte Programm ermöglicht außerdem eine "in place"-Berechnung, d.h. man braucht außer dem Platz für den reellen Datenvektor fast keinen zusätzlichen Speicherplatz. Bei der inversen reellen SRFFT erhält man aus $N/2 + 1$ Realteilen und $N/2 - 1$ Imaginärteilen einer konjugiert komplexen Funktion einen reellen Vektor aus N Datenwerten. Die Real- und Imaginärteile befinden sich vor der inversen SRFFT in einem reellen Vektor, in dem sich nach der Transformation das Ergebnis befindet; die Anordnung der Werte ist aber etwas ungewöhnlich mit

$$[\text{Re}(0), \text{Re}(1), \dots, \text{Re}(N/2), \text{Im}(N/2-1), \dots, \text{Im}(1)]$$

Zur Berechnung der Wahrscheinlichkeitsdichte aus der charakteristischen Funktion benötigen wir allerdings eine Vorwärtstransformation der Real- und Imaginärteile. Diese kann ohne Schwierigkeiten durch eine inverse Fouriertransformation ersetzt werden mit

$$\mathcal{F}\{f(u)\} = \mathcal{F}^{-1}\{f^*(u)\}, \quad (5.33)$$

also durch Änderung des Vorzeichens der Imaginärteile mit

$$[\operatorname{Re}(0), \operatorname{Re}(1), \dots, \operatorname{Re}(N/2), -\operatorname{Im}(N/2-1), \dots, -\operatorname{Im}(1)]$$

Nach einer Vertauschung der Intervalle für die positiven und negativen Werte von x in $p(x)$ mit Hilfe von Gleichung (5.18), nämlich

$$p(n + N/2 + 1) = p'(n) \quad \text{für } n = 0, 1, \dots, N/2 \quad (5.34)$$

und

$$p(n - [N/2 + 1]) = p'(n) \quad \text{für } n = N/2, N/2 + 1, \dots, N \quad (5.35)$$

haben wir eine Näherung für die Wahrscheinlichkeitsdichte am Ausgang des Empfängers.

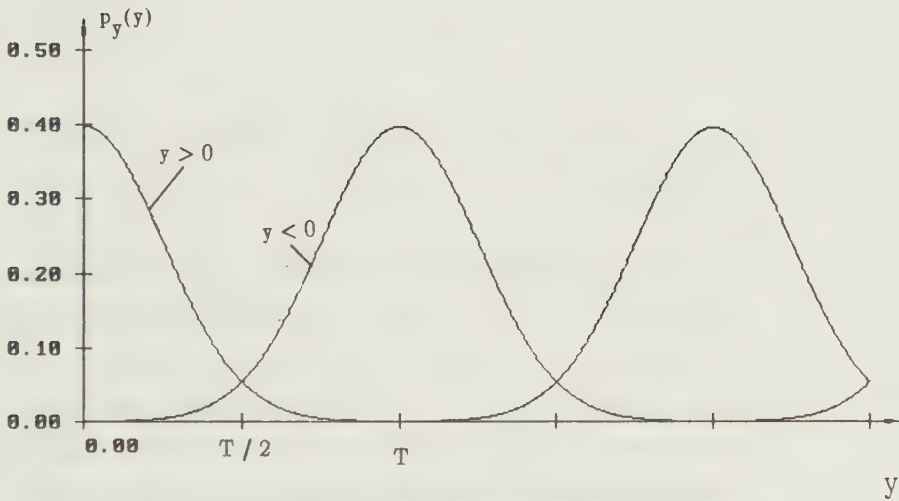


Bild 5.1.a Periodisierung der Wahrscheinlichkeitsdichte

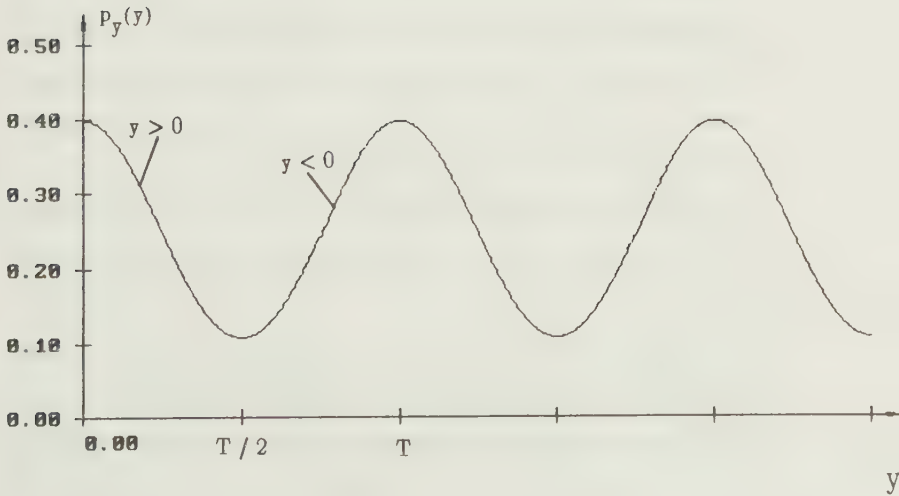


Bild 5.1.b Überlappung (Aliasing) der Wahrscheinlichkeitsdichte

6. Simulation

Um die im vorangegangenen Kapitel beschriebene numerische Methode zur Berechnung der Dichte zu untermauern, soll die Wahrscheinlichkeitsdichte am Ausgang des Empfängers zusätzlich mit Hilfe eines Simulationsverfahrens bestimmt werden. Wegen der stochastischen Eingangsgrößen (Rauschen und Signal) handelt es sich dabei um ein typisches Beispiel für die sogenannte Monte-Carlo-Simulation. Wichtigstes Merkmal solcher Simulationen ist die Benutzung von Zufallszahlenfolgen, die man als Realisation eines diskreten stochastischen Prozesses (hier: Rauschen) auffassen könnte.

Diese Methode, die von den Mathematikern Metropolis und Ulam 1949 [47] zum erstenmal als Monte-Carlo-Simulation bezeichnet wurde, erhielt ihren plastischen Namen von dem Verfahren, mit dem die Zufallszahlen erzeugt wurden: Aus den in Monte Carlo regelmäßig aushängenden Ergebnissen der Rouletteispiele wurden Tafeln von Zufallszahlen gewonnen.

Heute werden (Pseudo-)Zufallszahlenfolgen beinahe ausschließlich mit Hilfe von Digitalrechnern erzeugt; der Name Monte-Carlo-Simulation blieb aber erhalten.

Der Einsatz von Simulationen ist besonders dann sinnvoll, wenn es darum geht, schnell eine grob approximative Lösung des Problems zu erhalten, oder wenn andere numerische Lösungen nicht möglich zu sein scheinen oder zu

aufwendig sind. Dagegen sind Simulationsverfahren anderen numerischen Lösungen oft – wenn auch nicht immer – an Geschwindigkeit und Genauigkeit unterlegen, da sich bei strengen Forderungen an die Genauigkeit vor allem die langsame Konvergenz der Simulationsergebnisse negativ bemerkbar macht.

In unserem Fall soll eine Lösung sowohl numerisch als auch mit Hilfe eines Simulationsverfahrens erzielt werden. Der Sinn dieser Strategie liegt weniger in einem Vergleich der beiden Methoden; vielmehr wollen wir an Hand einiger Beispiele versuchen, die theoretischen Resultate mit Simulationsergebnissen zusätzlich zu untermauern. Da numerische- und Simulationsverfahren auf demselben Digitalrechner durchgeführt und graphisch ausgegeben werden können, besteht so die Möglichkeit, die Ergebnisse sofort und an Ort und Stelle zu verifizieren, wobei natürlich nie ganz ausgeschlossen werden kann, daß beide Verfahren den gleichen Fehler enthalten.

Die Zuverlässigkeit der Simulationsergebnisse hängt in unserem Fall vor allem davon ab, ob es gelingt, den untersuchten Empfänger (Bild 6.1)

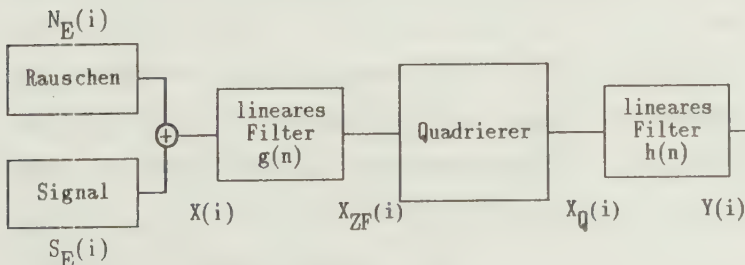


Bild 6.1 Der quadratische Detektor

auf geeignete Weise nachzubilden.

Zur Simulation der Rauschquelle muß, unter der Voraussetzung, daß es sich um thermisches Rauschen handelt, eine Folge von statistisch unabhängigen und normalverteilten Zufallszahlen erzeugt werden.

Die Erzeugung der Zufallszahlen ist dabei das größte Problem; die Umsetzung der beiden linearen Filter und des Quadrierers in ein Simulationsprogramm ist dagegen unproblematisch und soll nicht weiter behandelt werden, da es sich um digitale Baugruppen handelt.

Im letzten Schritt wird dann aus der Zahlenfolge am Ausgang des simulierten Empfängers die Wahrscheinlichkeitsdichte bestimmt.

6.1 Die Schätzung der Wahrscheinlichkeitsdichte

wird in Form eines Histogramms vorgenommen. Man erhält ein solches Histogramm, indem man den betrachteten Zahlenbereich in eine bestimmte Anzahl von Klassen unterteilt, und dann feststellt, wie viele Zahlen der Ausgangszahlenfolge in jede Klasse fallen. Wir teilen dazu das Intervall $(-SK/2, SK/2)$ in N gleichgroße Klassen $W(i)$ ein, wobei i eine ganze Zahl ist. Ist Y die Zahl, die die Ausgangsspannung des Empfängers simuliert, dann erhalten wir das Histogramm, indem wir das Intervall $(-SK/2, SK/2)$ durch Rundung auf die nächste ganze Zahl ($NINT$) auf die Klassen

$$W(i) \quad i = 1, 2, \dots, N \quad (6.1)$$

abbilden, mit

$$i = \text{NINT}(Y/SK \cdot N) + N/2 \quad (6.2)$$

Der Skalierungsfaktor SK hat dabei die gleiche Größe wie der Faktor, der für die Fouriertransformation der charakteristischen Funktion (Kapitel 5.6) verwandt wurde. Damit ist die Vergleichbarkeit der Ergebnisse gesichert.

Eine Schätzung für die Wahrscheinlichkeitsdichte von $p(x)$ erhält man nun durch Bestimmung der mittleren Häufigkeit für eine große Anzahl M von Simulationswerten Y.

$$P_s(i) = W(i) / M \quad (6.3)$$

Die Zahlen Y entstehen, wie bereits angedeutet, durch Filterung, Quadratur und eine weitere Filterung einer Summe aus einem deterministischen Signal und statistisch unabhängigen normalverteilten Abtastwerten.

6.2 Erzeugung normalverteilter Zufallszahlen

Normalverteilte Pseudozufallszahlen erhält man üblicherweise durch Transformation von statistisch unabhängigen sogenannten Standardzufallszahlen $U(0,1)$, die auf dem Intervall $(0,1)$ gleichverteilt sind. Verschiedene Methoden zur Erzeugung solcher Zahlen sollen zum Schluß vorgestellt werden.

Die im Programm benutzte Transformation von Standardzufallszahlen wurde von Box und Muller entwickelt und ist unter dem Namen "Polarmethode" bekannt. Sie beruht auf der Tatsache, daß Zufallsvektoren aus zwei aufeinanderfolgenden Zufallsvariablen in Polarkoordinaten einen in $(0, 2\pi)$ gleichverteilten Winkel besitzen und das Quadrat des Betrages exponentialverteilt ist. Die Erzeugung von exponentialverteilten Zufallszahlen ist aber im Gegensatz zu normalverteilten direkt möglich, indem man den Logarithmus einer Standardzufallszahl bildet. Man erhält dann jeweils zwei normalverteilte Zufallszahlen X_1, X_2 aus dem $U(0,1)$ verteilten Standardzufallszahlen Z_1, Z_2 , mit

$$\begin{aligned} X_1 &= \sqrt{-2 \cdot \ln(Z_1)} \cdot \sin(2\pi Z_2) \\ X_2 &= \sqrt{-2 \cdot \ln(Z_1)} \cdot \cos(2\pi Z_2) \end{aligned} \quad (6.4)$$

wobei X_1 und X_2 stochastisch unabhängige normalverteilte Zufallszahlen mit dem Mittelwert 0 und der Standardabweichung 1 sind. Das Verfahren setzt aber voraus, daß die Standardzufallszahlen gute statistische Eigenschaften haben. Verschiedene Methoden zur Erzeugung solcher Zahlen werden nun vorgestellt.

6.3 Erzeugung gleichverteilter Zufallszahlen

Anfangs wurden vor allem Tafeln mit Zufallszahlen benutzt, mit der Verbreitung der Digitalrechner wurden sie aber von anderen Methoden

verdrängt. Zur Erzeugung solcher Tafeln wurden neben den erwähnten Roulettspielen auch Daten einer Volkszählung, Telefonnummern und Logarithmentafeln verwendet ([19] S. 8). Die statistischen Eigenschaften der so erzeugten "Zufallszahlen" waren aber nicht immer befriedigend. Ein besonders kurioses Verfahren ist der aufwendige elektromechanische Erzeugungsmechanismus, mit dem M.G. Kendall und B. Babinton Smith 1940 eine Tafel aus 100 000 Ziffern generierten:

"Sie benutzten ein Gerät, bei dem eine Scheibe auf die Achse eines Elektromotors montiert war, der bis zu 2500 Umdrehungen pro Minute durchführen konnte. Die Scheibe war in 10 gleiche Segmente unterteilt, die mit den Ziffern 0 bis 9 beschriftet waren. Die rotierende Scheibe wurde in gewissen Zeitabständen durch einen Lichtblitz aus einer Neonlampe beleuchtet, so daß sie für das Auge eines Beobachters zu ruhen schienen. Dieser konnte nun die Nummer des Sektor notieren, auf den ein fest montierter Zeiger wies. Der Lichtblitz wurde durch den Entladungsstrom eines parallel zur Lampe geschalteten Kondensators hervorgerufen. Die Aufladung des Kondensators erfolgte über einen Widerstand, der aus einem Gewirr von Bleistiftlinien auf einem gewöhnlichen Stück Papier bestand, über das ein Kontaktstift glitt. Der Beobachter verschob mit der linken Hand ohne hinzusehen den Kontaktstift und notierte mit der rechten Hand die beobachtete Ziffer, wenn der Lichtblitz kam. Auf diese Weise wurden in vielen Sitzungen im Mittel 1500 Ziffern pro Stunde produziert. Um kleinen

Unregelmäßigkeiten in der Scheibe entgegenzuwirken, wurde diese nach jeder Sitzung gegen eine andere, gleichartige Scheibe ausgetauscht." (zitiert nach [53], S. 14)

Zur Benutzung auf Digitalrechnern kann man solche Tafeln auf Massenspeichern wie Magnetbändern oder Platten ablegen und wieder einlesen, wenn sie benötigt werden. Nachteil dieses Verfahrens ist neben dem Speicherplatzbedarf vor allem, daß der Zugriff auf Massenspeicher auf allen Rechnern zu den zeitaufwendigsten Operationen gehört, die Benutzung von Zufallstafeln wird damit langsam und teuer.

Andere Quellen "echter" Zufallszahlen, sogenannte physikalische Generatoren, haben sich ebenfalls nicht durchsetzen können. Hier werden Rauschvorgänge oder der Zerfall radioaktiver Substanzen zur Erzeugung zufälliger Zahlenreihen benutzt.

Beobachtet man beispielsweise die Spannung des thermischen Rauschens an einem Widerstand, so kann man, wenn der zeitliche Abstand zwischen den Beobachtungen genügend groß gewählt wurde, von unkorrelierten und damit auch statistisch unabhängigen Abtastwerten (Gauß-Prozeß) ausgehen. Wählt man nun eine Schwelle c , so daß für

$$X_i = \begin{cases} 1, & \text{wenn } U(t_i) \geq c \\ 0, & \text{wenn } U(t_i) < c \end{cases} \quad (6.5)$$

gilt

$$P(X_i = 1) = P(X_i = 0) = 0.5 \quad (6.6)$$

so erhält man damit eine Folge von statistisch unabhängigen binären Ziffern.

Jeweils k solcher Ziffern können dann zu einer k -stelligen Binärzahl zusammengefaßt werden.

Die Nachteile dieses Generators sind typisch für die gesamte Klasse der physikalischen Verfahren: Da die Schwelle c sehr genau eingehalten werden muß, sind Wartungsarbeiten unvermeidlich, und Umwelteinflüsse wie zum Beispiel Temperaturschwankungen oder Hochfrequenzeinstrahlungen können die statistischen Eigenschaften extrem verschlechtern. Außerdem wirkt sich die Bedingung, daß der Abstand zwischen den Abtastzeitpunkten nicht beliebig klein gewählt werden kann und jeweils k Beobachtungen für eine Zufallszahl benötigt werden, negativ auf die Geschwindigkeit des Generators aus.

Heute werden beinahe ausschließlich sogenannte Pseudozufallszahlengeneratoren verwendet. Es handelt sich dabei um arithmetische Methoden, die zu einer streng determinierten und sogar periodischen Zahlenfolge führen, also strenggenommen nichts weniger als zufällig sind.

John von Neumann, der selbst einen arithmetischen Zufallszahlengenerator entwickelt hat, faßte dies in einem, inzwischen berühmt gewordenen Satz so zusammen:

"Anyone who considers arithmetical methods of producing random digits is, of course, in state of sin."

(Zitiert nach D.E. [39], S.1) Trotzdem sind solche Zahlenreihen so "chaotisch", daß –im Vergleich zur Periodenlänge– kurze Sequenzen mit statistischen Mitteln nicht oder nur sehr schwer von Realisationen echter Zufallsvariablen unterschieden werden können. Solche Methoden sind bereits

im Einsatz, seit es Digitalcomputer gibt, zu ihren Vorteilen zählen vor allem geringer Zeit- und Speicherplatzbedarf. Die Zweifel an diesen Methoden sind allerdings genauso alt wie die Verfahren selbst. Neben grundsätzlichen Bedenken (v. Neumann) und eher psychologisch motivierten Zweifeln, die die Einfachheit betreffen, sind aber auch konkrete Mängel an den erzeugten Zahlenreihen festgestellt worden, wie wir noch sehen werden.

6.4 Lineare Kongruenzgeneratoren

Der einfachste Vertreter dieser Methode benötigt zwar am wenigsten Rechenzeit, hat aber auch die schlechtesten "statistischen" Eigenschaften. Das Verfahren heißt linearer Kongruenzgenerator und ist definiert durch die Anweisung ([53] S. 24)

$$y_{i+1} = a \cdot y_i + c \pmod{m} \quad (6.7)$$

Dabei sind die Konstanten

m (der Modul) a (der Faktor)
 c (das Inkrement) Y_0 (die Startzahl)
 ganze Zahlen mit $0 \leq a, c, y_0$

Die rationalen Zahlen

$$z_i = y_i / m \quad (6.8)$$

verwendet man dann als Pseudozufallszahlen.

Die Operation $\pmod{m}$ bedeutet dabei, daß der Rest einer ganzzahligen

Division durch m als Ergebnis zu nehmen ist; z.B. ist

$$k \cdot m + \ell \pmod{m} = \ell \quad \text{für } \ell < m. \quad (6.9)$$

Da jede ganze Zahl in die Form (6.9) überführt werden kann, ergibt sich aus Gleichung (6.9), daß die Operation $\pmod{m}$ jede positive ganze Zahl in das Intervall $[0, m)$ abbildet.

Diese Eigenschaft ist der Grund für die "Zufälligkeit" der Zahlenfolge.

Eine weitere Konsequenz der Modulo-Operation ist, daß sich die Zahlenfolge mit einer Periode von höchstens m wiederholt. Bei schlecht gewähltem Faktor a kann sie aber auch wesentlich kleiner als m sein. Der Faktor a beeinflußt außerdem die "statistischen" Eigenschaften der Zahlenfolge z_i .

Für $c = 0$ spricht man auch von multiplikativen Kongruenzgeneratoren, ist $c \neq 0$, hat man einen gemischten Kongruenzgenerator.

Die Geschichte dieser Generatoren, die auch als Lehmer-Generatoren bezeichnet werden, hat Marsaglia ([46] S.276) so zusammengefaßt:

"Congruential generators were apparently first suggested as random number generators in 1948 by D.H. Lehmer,...In the Period 1948-1968, linear congruential sequences were by far the most commonly used method for generating uniform random variables in computers, and most users were satisfied with them. Hundreds of papers and reports were written on choices of multiplier, additive constant and modulus, and giving results of tests for uniformity. Most of the tests were satisfactory, although there were a few examples of generators that gave unsatisfactory results."

Diese Situation änderte sich 1968 grundlegend, als Marsaglia [45] beweisen konnte, daß *alle* linearen Kongruenzgeneratoren die folgende bemerkenswerte Schwäche aufweisen:

Bildet man n -Tupel aus n aufeinander folgenden Zahlen der Pseudo-zufallszahlenfolge, so liegen alle so gewonnenen Punkte auf einer gewissen Anzahl von parallelen Hyperebenen im n -dimensionalen Raum. Die maximal mögliche Zahl der Ebenen ist, wie Marsaglia gezeigt hat, $(n!m)^{1/n}$, sie hängt also vom verwendeten Modul ab.

Daraus ergibt sich, daß bei $m = 2^{31}$ (diesen Modul findet man bei 32-Bit Rechnern am häufigsten) jeweils drei aufeinander folgende Zahlen auf maximal 2344 Ebenen und 10 aufeinanderfolgende Zahlen auf höchstens 38 parallelen Ebenen liegen können. Bei ungünstig gewähltem Multiplikator kann die Anzahl der Ebenen aber auch erheblich niedriger sein. Ein bekanntes Beispiel ist der 1970 von IBM als Teil des "System/360 scientific subroutine package" [62] veröffentlichte Generator RANDU.

Allein die Nennung des Namens RANDU soll nach D.E. Knuth ausreichen, um bei vielen Computer - Wissenschaftlern Bestürzung auszulösen ([39], S.28).

Bildet man aus der Zahlenfolge dieses Generators Tripel aus aufeinanderfolgenden Zahlen,

$$u_i = (u_{3i-2}, u_{3i-1}, u_{3i}) \quad i = 1, 2, 3, \dots, \quad (6.10)$$

so liegen diese auf nur 15 parallelen Ebenen im dreidimensionalen Raum.

Trotz dieser bekannten Eigenschaft ist RANDU nach Einschätzung von

Dudewicz und Ralley [19] der wahrscheinlich weltweit am stärksten verbreitete Zufallszahlengenerator. Nach eigenen Erfahrungen stellt die Universität Bremen dabei keine Ausnahme dar. So ist er der einzige Generator, der auf dem Großrechner des Regionalen Rechenzentrums fest implementiert ist.

Bild 6.2 zeigt die dreidimensionale Verteilung der generierten Zahlen im Vergleich zu der Verteilung, die der im Simulationsprogramm benutzten Generator URN12 erzeugt (6.3). Beide Generatoren unterscheiden sich nur durch den Faktor a . RANDU ist definiert durch:

$$y_{i+1} = 65539 \cdot y_i \pmod{2^{31}}, \quad (6.11)$$

während die Zahlenfolge von URN12 [19] erzeugt wird mit

$$y_{i+1} = 452807053 \cdot y_i \pmod{2^{31}} \quad (6.12)$$

Aus Bild 6.3 darf aber nicht geschlossen werden, daß die Zahlenfolge des Generators URN12 nicht auf parallelen Ebenen liegt. Die Anzahl der Ebenen ist nur zu groß um sie noch graphisch darstellen zu können. Ihre genaue Zahl ist leider unbekannt; sie muß aber wegen des Moduls von 2^{31} kleiner sein als 2344. Der Generator wurde jedoch von Dudewicz und Ralley [19] 20 verschiedenen statistischen Tests unterzogen, ohne daß sich Mängel ergeben hätten.

Seit 1968 sind große Anstrengungen unternommen worden, um Pseudozufallszahlengeneratoren zu entwickeln, die eine hochdimensionale Gleichverteilung

ohne oder mit sehr vielen parallelen Ebenen besitzen. Einige der Veröffentlichungen zu diesem Thema scheinen jedoch recht unseriös zu sein. So schlägt John R.B. Whittlesey 1969 [64] einen Schieberegister-Zufallsgenerator (siehe Kapitel 2.2) als Alternative zu Kongruenzgeneratoren vor, da Schieberegistergeneratoren angeblich eine nahezu perfekte n -dimensionale Gleichverteilung erzeugen können.

G. Marsaglia bestreitet die in diesem Aufsatz vertretene These jedoch entschieden [46]. Seiner Meinung nach sind Schieberegistergeneratoren noch schlechter als Kongruenzgeneratoren und besitzen Regelmäßigkeiten für höhere Dimensionen, die denen von Kongruenzgeneratoren ähnlich sind. Einen Beweis dafür findet man bei Marsaglia [46] aber nicht.

Tatsächlich fallen Schieberegistergeneratoren bei vielen statistischen Tests durch, die von Kongruenzgeneratoren bestanden werden (Dudewicz und Ralley 1980).

Eigene Untersuchungen haben gezeigt, daß eine verallgemeinerte Version des Generators von Whittlessy [40] nicht einmal unkorrelierte Zahlen erzeugt, da das Leistungsdichtespektrum ungleichförmig ist (Bild 6.5, Bild 6.6 zeigt das LDS des Generators URN12). Der Generator kann dann auch keine zwei- oder dreidimensional gleichverteilten Zahlen erzeugen, wie von Lewis und Payne [40] behauptet wurde. Die dreidimensionale Verteilung haben Lewis und Payne mit Hilfe eines graphischen Verfahrens (wie in Bild 6.2) untersucht. Solche Methoden führen aber auch bei Kongruenzgeneratoren zu guten Testergebnissen, wenn die Anzahl der Ebenen nicht allzu klein ist. Man geht deshalb immer mehr von statistischen auf geometrische Tests über, was bedeutet, daß man nicht mehr die Zahlenfolge der Generatoren

untersucht, sondern die Lage der Ebenen mit Hilfe von zahlentheoretischen Methoden direkt aus der Generatorfunktion bestimmt. Solche Methoden sind für Kongruenzgeneratoren bereits weit entwickelt. Für Schieberegistergeneratoren scheint dies aber mit großen Schwierigkeiten verbunden zu sein, was ihre Weiterentwicklung behindert.

6.5 Mehrfach rekursive Generatoren

Diese Verallgemeinerung des linearen Kongruenzverfahrens ist 1969 von D.E Knuth ([39], S. 28) vorgeschlagen worden

$$y_i = a_1 \cdot y_{i-1} + a_2 \cdot y_{i-2} + \dots + a_k \cdot y_{i-k} \pmod{m} \quad (6.13)$$

Dabei sind $a_1, a_2, \dots, a_k$ die Faktoren, m ist der Modul und $y_{i-1}, y_{i-2}, \dots, y_{i-k}$ ist der Startvektor. $U(0,1)$ - verteilte Zahlen erhält man mit $z_i = y_i/m$.

Dieses Verfahren nennt man einen mehrfach rekursiven Generator. Die Periode eines k -fach rekursiven Generators kann bei gut gewählten Faktoren m^{k-1} betragen, während gewöhnliche Kongruenzgeneratoren eine maximale Periode von $m-1$ haben.

Bildet man bei k -fach rekursiven Generatoren n -Tupel aus n aufeinanderfolgenden Zufallszahlen, so liegen diese auch hier auf parallelen Ebenen, deren Anzahl allerdings mit

$$(n!m^k)^{1/n} \quad (6.14)$$

erheblich höher sein kann als bei (gewöhnlichen) einfach-rekursiven Generatoren. (A. Grube [29], S.1-5). Grube hat gezeigt, daß das Verhalten eines $(k+1)$ -fach rekursiven Generators im $(n+1)$ dimensionalen Raum dem eines k -fachen rekursiven Generators im n -dimensionalen Raum entspricht (A. Grube [29] S. 8.9). Die "statistischen" Eigenschaften werden deshalb mit steigendem k immer besser, wenn die Faktoren a_i gut gewählt sind. Die Rechenzeit pro Zahl ist aber mehr als k -fach so groß wie bei einfach rekursiven Kongruenzgeneratoren.

6.6 Nichtlineare Kongruenzgeneratoren

Grundsätzlich vermeiden kann man die parallelen Ebenen durch Verwendung eines speziellen nichtlinearen Kongruenzgenerators, den J. Eichauer und J. Lehn [24] vorgeschlagen haben. Die Methode ist definiert durch

$$y_i = a \cdot (y_{i-1})^{-1} \pmod{m} \quad (6.15)$$

Dabei ist m eine Primzahl, $a, b < m$ sind positive Konstanten, und y^{-1} ist eine ganzzahlige Inverse mit $0 < y^{-1} < m$, die durch

$$y^{-1} \cdot x \pmod{m} = 1 \quad (6.16)$$

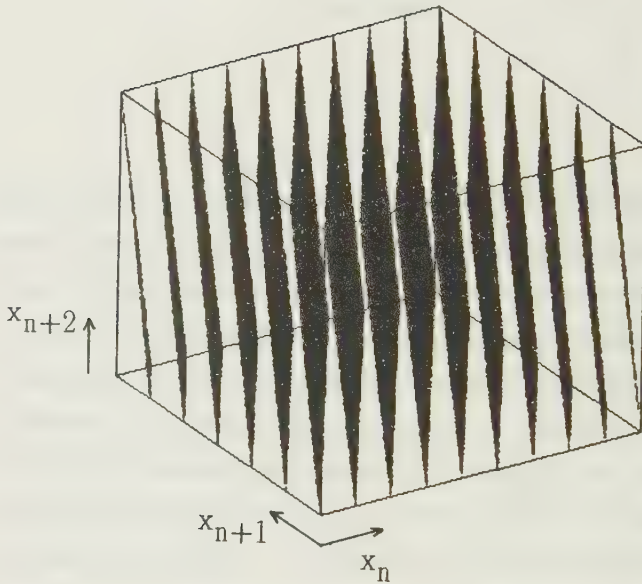
eindeutig definiert ist.

Wie J. Lehn in einem Vortrag erläuterte, den er 1987 an der Universität Bremen gehalten hat, sorgt die Inversion dafür, daß die parallelen Ebenen linearer Kongruenzgeneratoren zu Hyperbeln "verbogen" werden, wodurch sich die statistischen Eigenschaften erheblich verbessern sollen.

Allerdings kostet die Inversion erhebliche Rechenzeit. Bei einem Modul von $m \approx 10^6$ und dem von Eichauer und Lehn benutzten Algorithmus zur Berechnung von y^{-1} benötigt der Generator im Mittel zur Erzeugung einer Zufallszahl etwa zwölfmal so lange wie ein linearer Kongruenzgenerator. Die Anzahl der Hyperbeln läßt sich außerdem durch mehrfachrekursive Verfahren steigern. Ein mehrfach rekursiver nichtlinearer Kongruenzgenerator wurde 1987 von Eichauer, Grothe, Lehn und Topuzoğlu [24] vorgeschlagen. Die Berechnung der Zahlen ist allerdings noch wesentlich komplizierter und zeitaufwendiger als bei dem zuvor beschriebenen Generator.

Zu den statistischen Eigenschaften nichtlinearer Kongruenzgeneratoren äußern sich die Autoren allerdings betont vorsichtig:

"The results of these empirical tests give no reason to dissuade one from using a non-linear congruential generator ... to generate a sequence of pseudo random numbers." [23].



Zufallszahl
generator

$x = ax + c \pmod{m}$

$a=65539$

$c=0$

$m=2147483647$

$\alpha=33^\circ$ $\rho=6$

Bild 6.2.a. Dreidimensionale Verteilung des Generators RANDU.

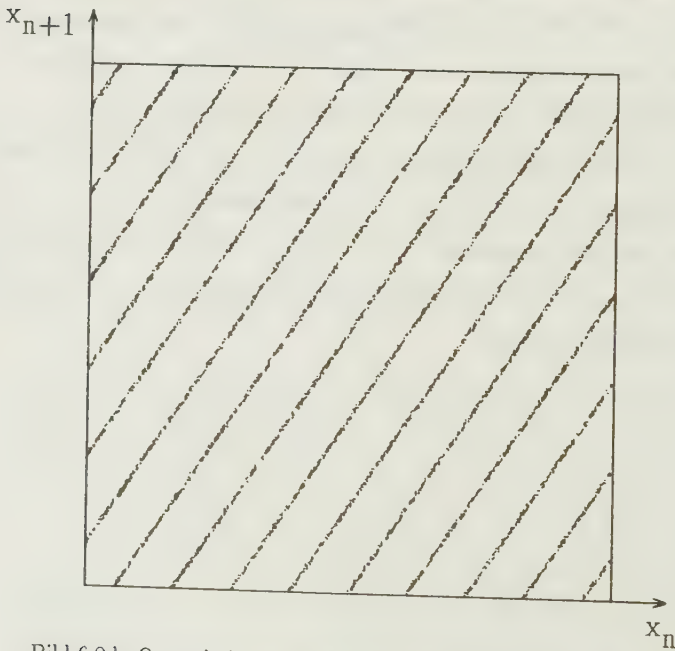
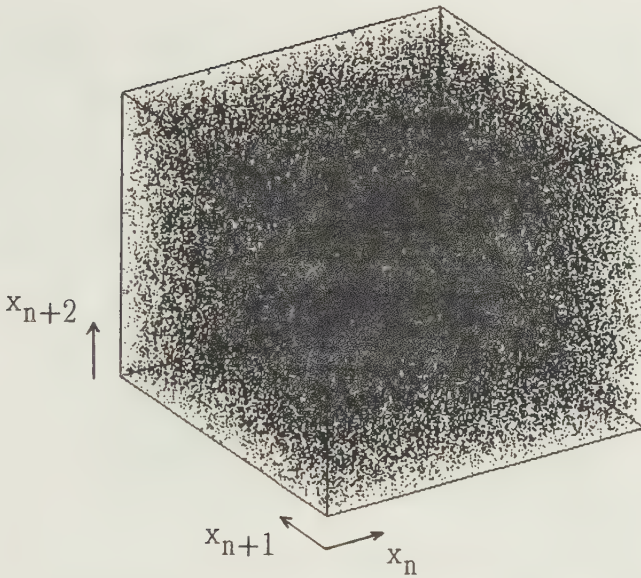


Bild 6.2.b. Querschnitt durch die dreidimensionale Verteilung
bei $x_{n+2} = 0,5$.



**Zufallszahlen-
generator:**

$x = ax + c \pmod{m}$

$a=452807053$

$c=0$

$m=2147483648$

$\alpha=33^\circ \quad \beta=60^\circ$

Bild 6.3.a. Dreidimensionale Verteilung des Generators URN12.

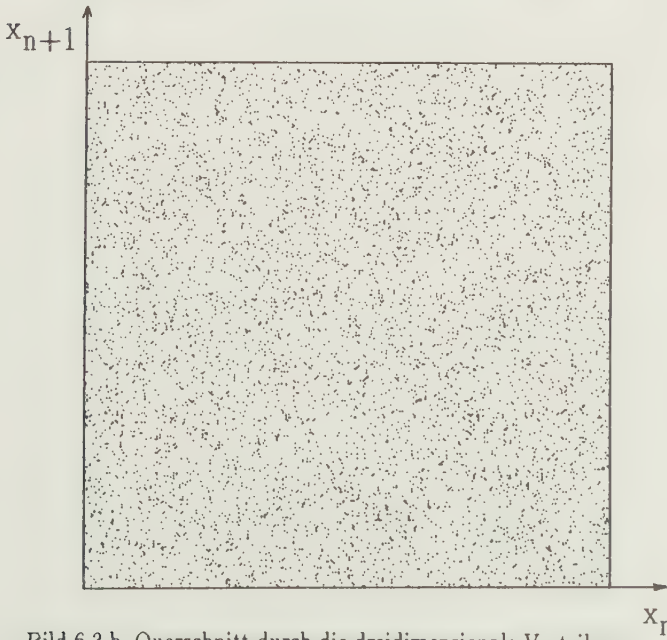
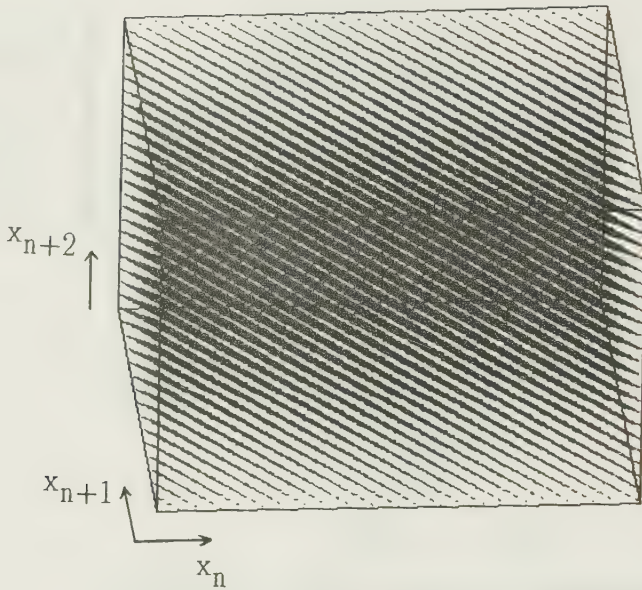


Bild 6.3.b. Querschnitt durch die dreidimensionale Verteilung

bei $x_{n+2} = 0,5$.



Zufallszahl
generator

$x = ax + c \pmod{m}$
 $a=7141$
 $c=54773$
 $m=259200$

$\alpha=5^\circ$ $\beta=5^\circ$

Bild 6.4.a. Dreidimensionale Verteilung eines gemischten
Kongruenzgenerators.

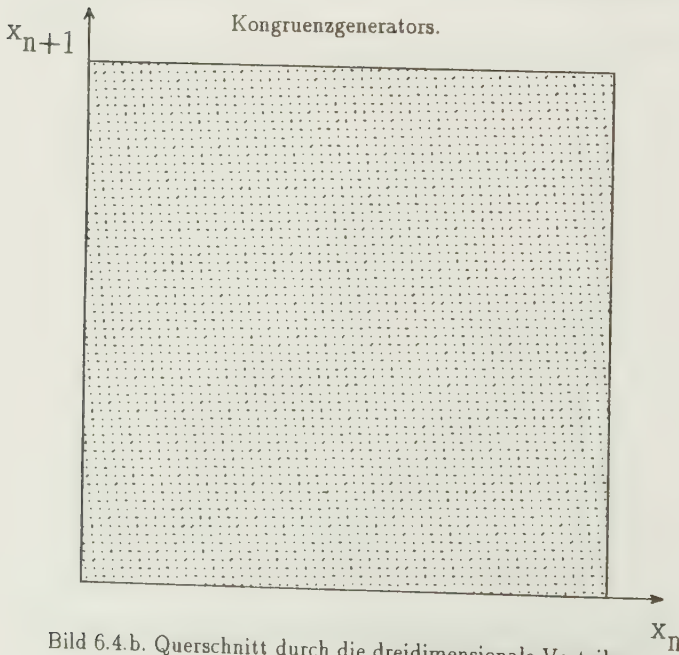


Bild 6.4.b. Querschnitt durch die dreidimensionale Verteilung
bei $x_{n+2} = 0,5$.

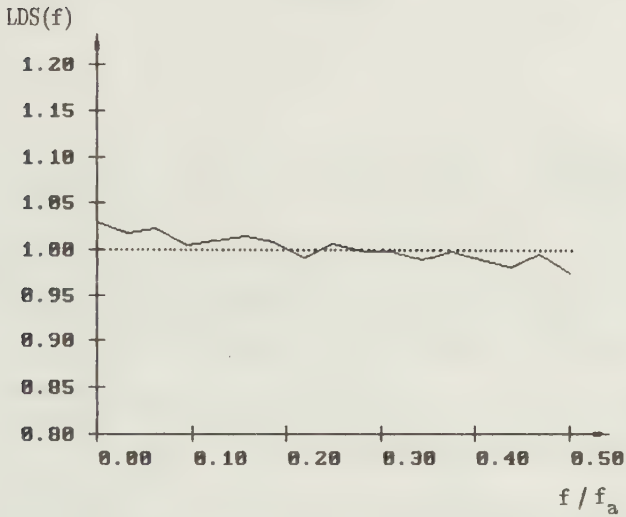


Bild 6.5.a. Leistungsdichtespektrum des Generators GFSR.

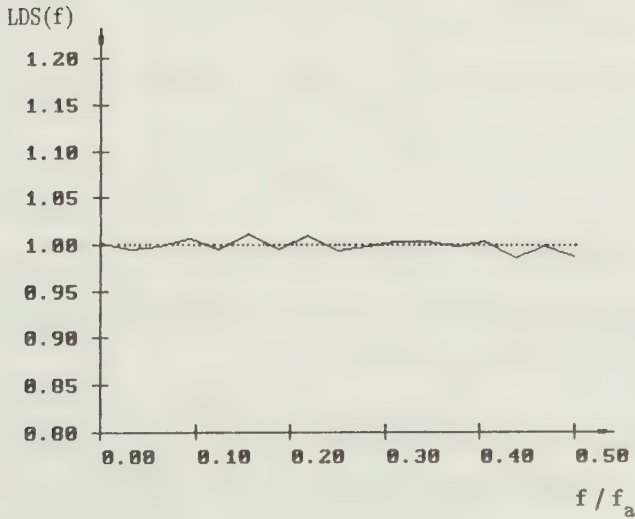


Bild 6.5.b. Leistungsdichtespektrum des Generators URN12.

7. Ergebnisse

Beim Empfang gestörter Signale treten unvermeidbare Fehler auf. Will man beispielsweise detektieren, ob zu einem bestimmten Zeitpunkt gesendet wird oder nicht, so treten in Abhängigkeit von der gewählten Entscheidungsschwelle Fehler auf, deren Wahrscheinlichkeit durch geeignete Wahl der Schwelle minimiert werden kann.

Zur Festlegung der Schwelle und zur Berechnung der Fehlerwahrscheinlichkeit benötigt man die gemeinsame eindimensionale Wahrscheinlichkeitsdichte des Signals und der Störung (Rauschen).

Die Bilder 7.1 bis 7.6 zeigen die charakteristische Funktion beziehungsweise die Wahrscheinlichkeitsdichte am Ausgang des in Kapitel 4 beschriebenen quadratischen Detektors (Bild 4.1).

Die eindimensionale charakteristische Funktion am Ausgang des Empfängers ist nach Gleichung 4.47

$$f_{NF_i}^k(u) = |\hat{I} - 2 j u \hat{H}_D \hat{A}|^{-1/2} \cdot \exp \{ j u \hat{S}_{ZF}^T (\hat{I} - 2 j u \hat{H}_D \hat{A})^{-1} \hat{H}_D \hat{S}_{ZF} \}$$

Aus der charakteristischen Funktion wird dann mit Hilfe der schnellen Fouriertransformation (Kapitel 5.6 und 5.7) die Wahrscheinlichkeitsdichte am Ausgang des Empfängers berechnet.

Das Sendesignal ist in allen Fällen $\sin(2\pi \cdot f \cdot i)$ mit der normierten Frequenz $f = 0.35$. Die in Kapitel 5 skizzierte Methode ermöglicht die Berechnung der Wahrscheinlichkeitsdichte ohne Einschränkung auch für andere Signale.

Bei den linearen Filtern $g(n)$ und $h(n)$ wurden nichtrekursive Kaiserfilter

verwendet. Das erarbeitete Programm ermöglicht die Berechnung der Wahrscheinlichkeitsdichte für beliebige stabile, rekursive oder nicht-rekursive Filter, wenn die Impulsantwort bekannt ist.

Für die Bilder 7.1 bis 7.6 wurden identische Bandpässe verwendet. Die Grenzfrequenzen sind $f_u = 0,15$ und $f_o = 0,45$. Die Filtersteilheit ist 0,05 , und die Mindestsperrdämpfung beträgt 30 dB.

Die Filtersteilheit der Tiefpässe $h(n)$ ist in allen Fällen ebenfalls 0,05, die Mindestsperrdämpfung 30 dB.

Die Bilder 7.1 bis 7.3 zeigen die charakteristische Funktion für verschiedene Grenzfrequenzen des NF-Filters $h(n)$, und zwar 0,15 in Bild 7.1.a und 7.1.b , 0,075 in Bild 7.2.a und 7.2.b und 0,0375 in den Bildern 7.3.a und 7.3.b.

Die Bilder 7.4 bis 7.6 zeigen die entsprechenden Wahrscheinlichkeitsdichten für die verschiedenen Bandbreiten, und zwar mit und ohne Sendesignal.

7.4.a , 7.5.a und 7.6.a wurden mit dem in Kapitel 5 beschriebenen numerischen Verfahren ermittelt, während die in Bild 7.4.b , 7.5.b und 7.6.b gezeigten Wahrscheinlichkeitsdichten mit Hilfe einer Monte-Carlo-Simulation gewonnen wurden.

Zur Simulation wurde der Generator URN12 (Gleichung 6.12) mit der Startzahl 1 benutzt. Die Ungenauigkeit des Simulationsergebnisses liegt in einer Größenordnung, die bei einer Simulation mit 20000 "Zufallszahlen" zu erwarten ist.

Das Simulationsergebnis bestätigt offenbar die theoretischen Überlegungen.

Man erkennt, daß in allen Fällen die Varianz der Ausgangsverteilung bei gesendetem Träger größer ist als ohne (Sendeleistung = 0). Dies ist mit der für große Signale zunehmenden Steigung der quadratischen Kennlinie zu

erklären. Man erkennt außerdem, daß die Varianz am Ausgang des Empfängers mit geringer werdender Bandbreite des NF-Filters abnimmt.

Ist ein Sendesignal vorhanden, nähert sich die Verteilung einer Normalverteilung mit dem Mittelwert 2.

Weiterhin sieht man, daß sich die Verteilungen (mit und ohne Sendesignal) mit abnehmender NF-Bandbreite immer weniger überlappen. Daraus folgt, daß die Wahrscheinlichkeit für eine falsche Detektion mit kleiner werdender Bandbreite abnimmt.

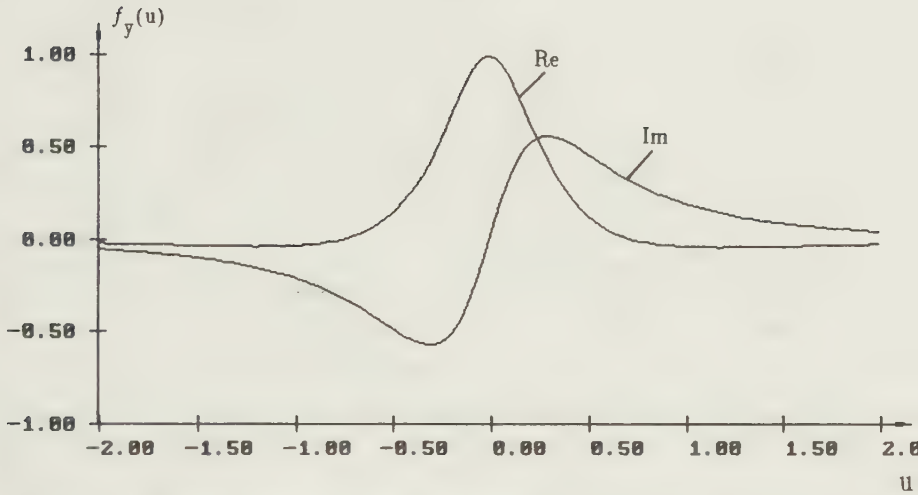


Bild 7.1.a. Charakteristische Funktion am Ausgang des Empfängers.
Die Bandbreite des NF - Filters beträgt 0,15
Die Sendeleistung ist 0.

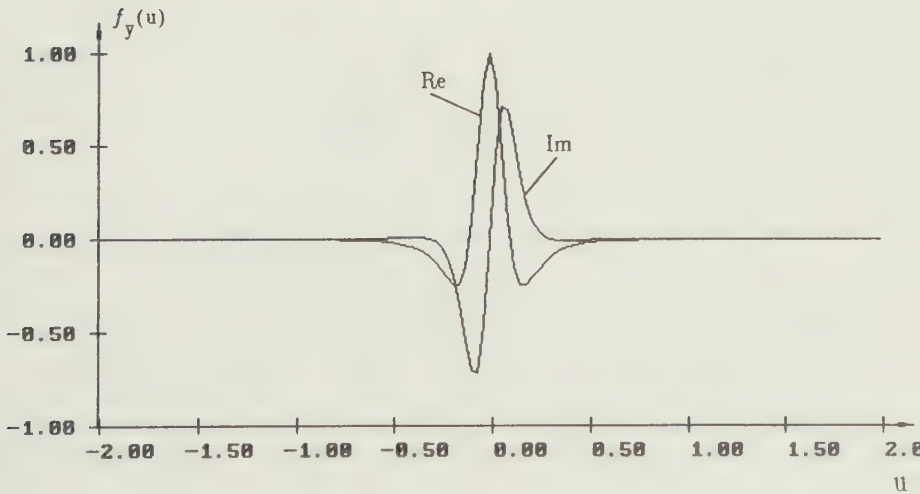


Bild 7.1.b. Charakteristische Funktion am Ausgang des Empfängers.
Die Bandbreite des NF - Filters beträgt 0,15
die Sendeleistung ist 2.

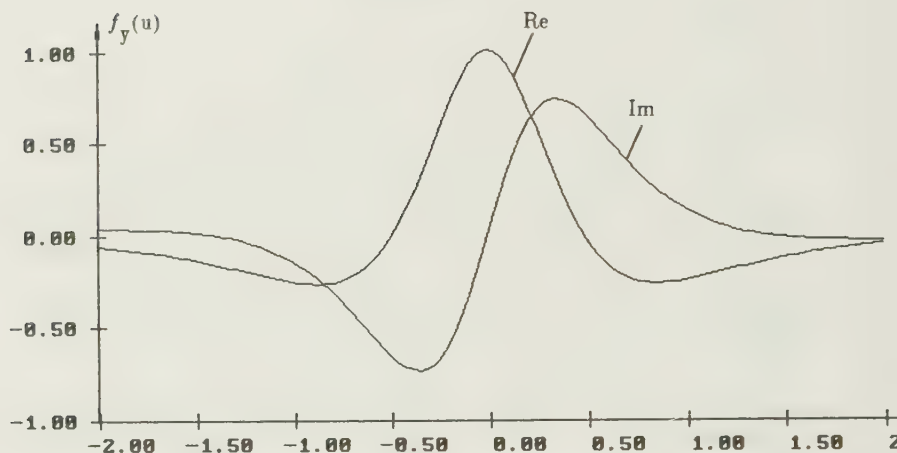


Bild 7.2.a. Charakteristische Funktion am Ausgang des Empfängers.
Die Bandbreite des NF - Filters beträgt 0.075
Die Sendeleistung ist 0.

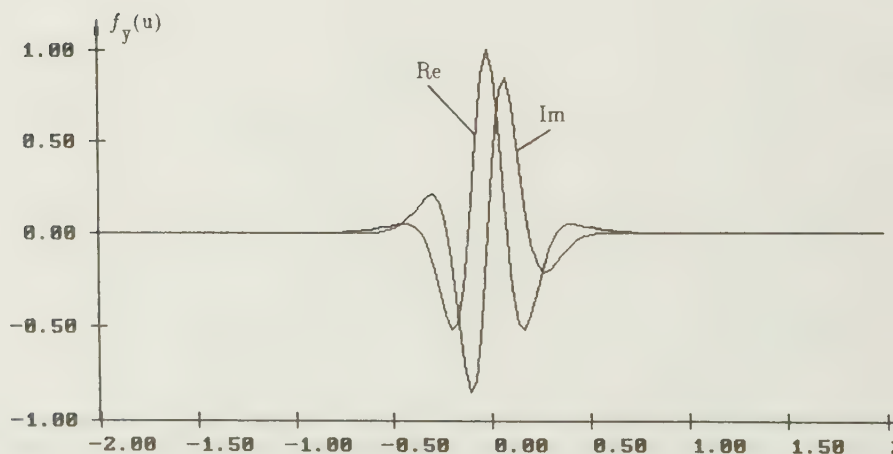


Bild 7.2.b. Charakteristische Funktion am Ausgang des Empfängers.
Die Bandbreite des NF - Filters beträgt 0,075
Die Sendeleistung ist 2.

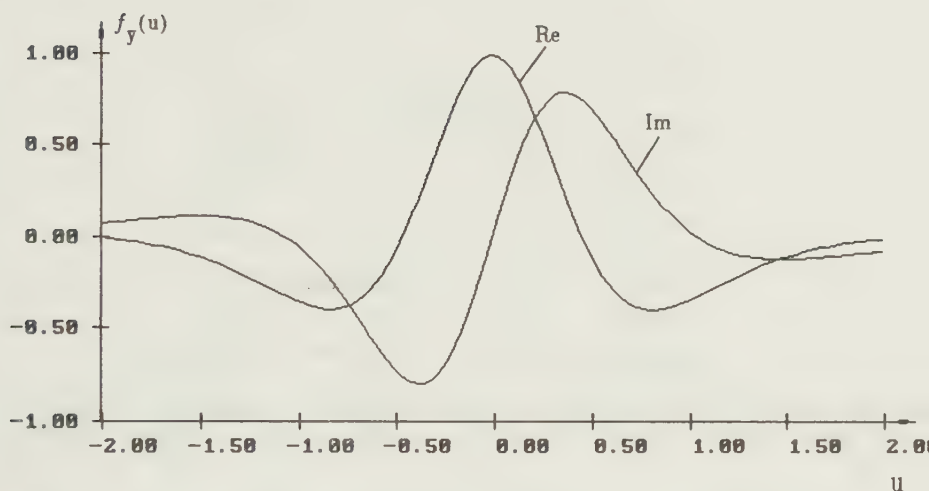


Bild 7.3.a. Charakteristische Funktion am Ausgang des Empfängers.
Die Bandbreite des NF - Filters beträgt 0,0375
Die Sendeleistung ist 0.

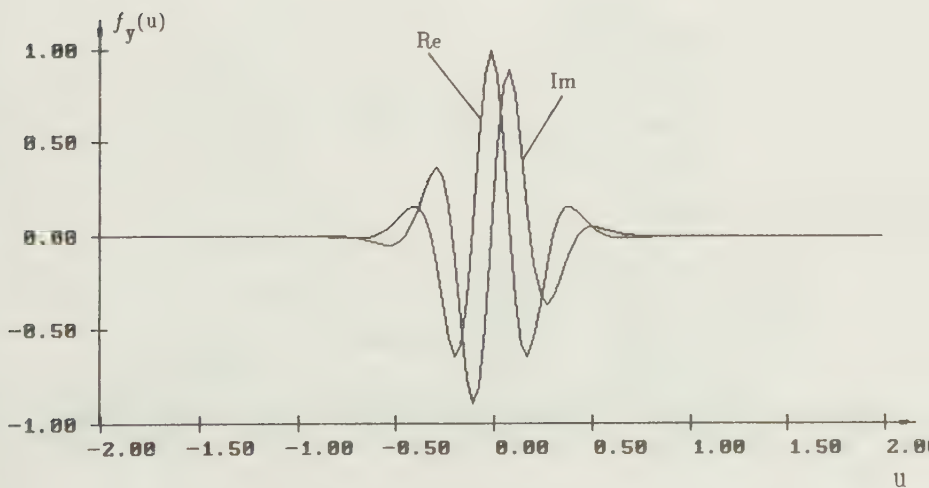


Bild 7.3.b. Charakteristische Funktion am Ausgang des Empfängers.
Die Bandbreite des NF - Filters beträgt 0,0375
Die Sendeleistung ist 2.

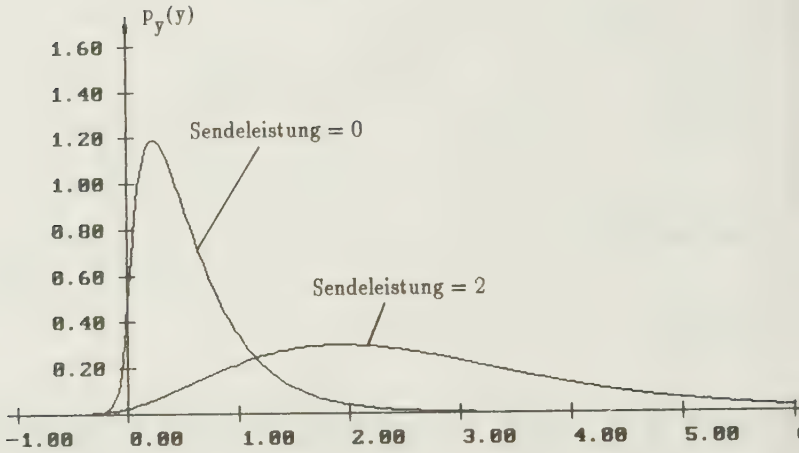


Bild 7.4.a. Wahrscheinlichkeitsdichte am Ausgang des Empfängers.

Die Bandbreite des NF – Filters beträgt 0,15.

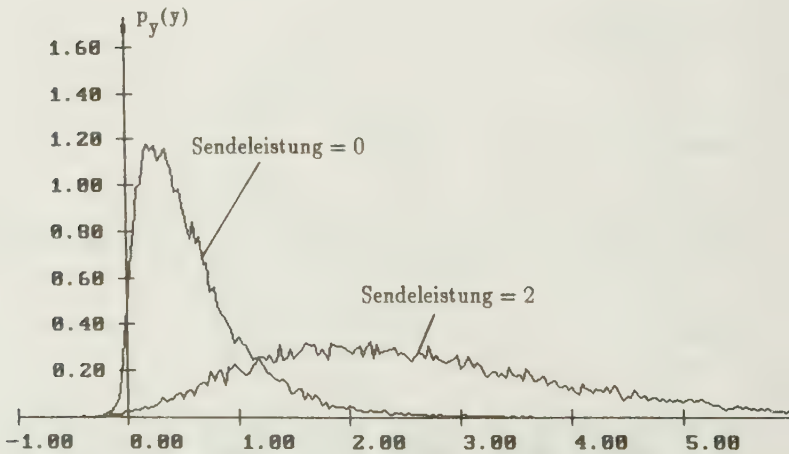


Bild 7.4.b. Ergebnis der Simulation

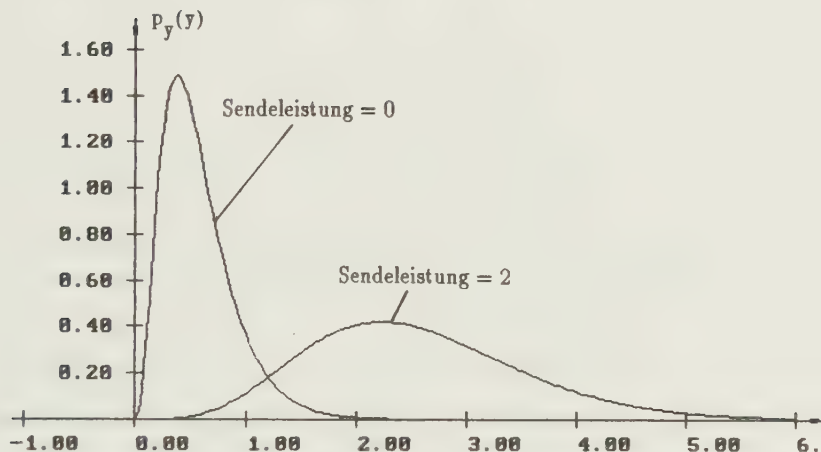


Bild 7.5.a. Wahrscheinlichkeitsdichte am Ausgang des Empfängers.

Die Bandbreite des NF – Filters beträgt 0,075.

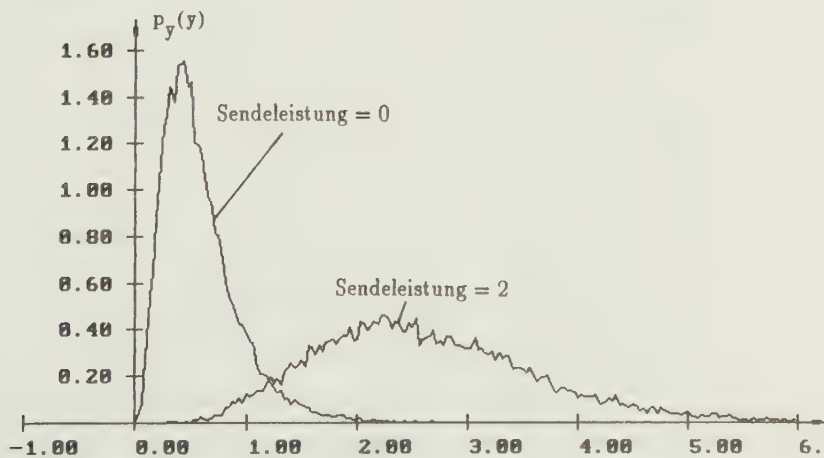


Bild 7.5.b. Ergebnis der Simulation

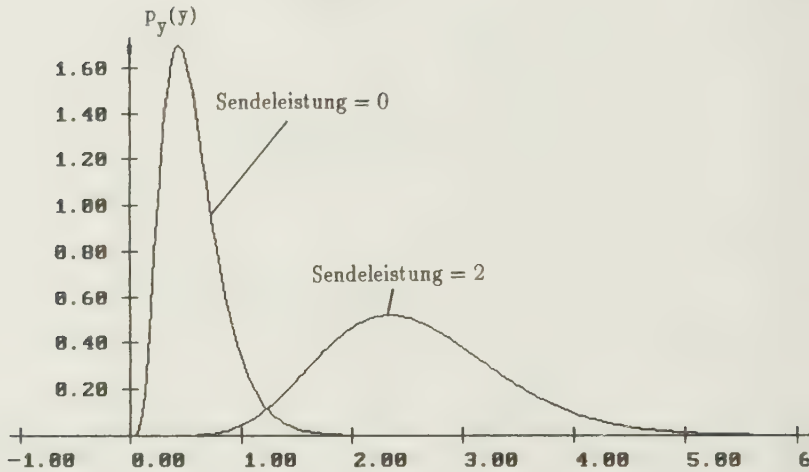


Bild 7.6.a. Wahrscheinlichkeitsdichte am Ausgang des Empfängers.

Die Bandbreite des NF – Filters beträgt 0,0375.

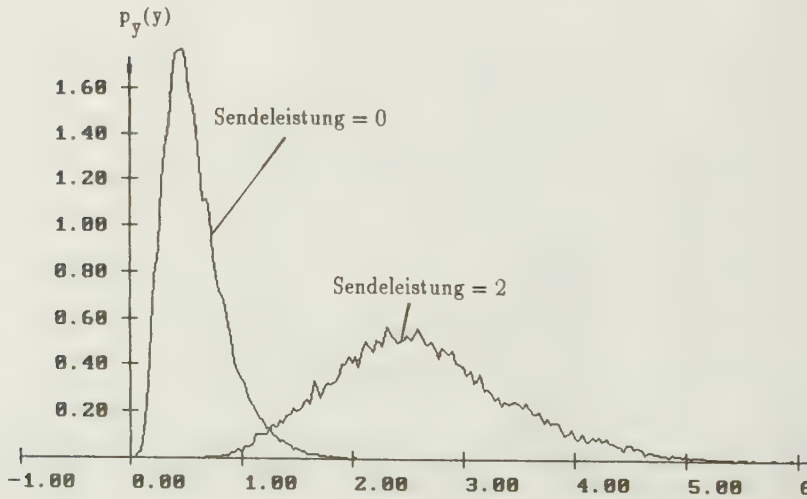


Bild 7.6.b. Ergebnis der Simulation

8. Zusammenfassung

Wie sich die Wahrscheinlichkeitsdichte eines stochastischen Prozesses unter dem Einfluß eines linearen Filters verändert, ist eine naheliegende Frage, die z.B. für die Berechnung der Fehlerwahrscheinlichkeit in Detektorsystemen von Bedeutung ist. In der Literatur wird in Zusammenhang mit linearen Filtern meist nur die Berechnung der AKF erwähnt (Wiener – Lee – Formel).

Bei analogen Filtern sind Lösungen für die Verteilung nur in wenigen Spezialfällen bekannt; z.B. für einen Detektor aus einem ZF-Bandfilter, einem Quadrierer und einem Tiefpaß und weißem Rauschen als Eingangssignal.

Ziel der vorliegenden Arbeit war es, einen möglichst allgemeingültigen Ansatz zur Bestimmung der Wahrscheinlichkeitsdichte von Rauschvorgängen nach digitaler Filterung zu entwickeln. Eine einfache Lösung für die Verteilung nach der Filterung erhält man nur für den Fall statistisch unabhängiger Abtastwerte und zwar mit Hilfe des Faltungssatzes der Statistik.

Läßt man beliebige statistische Abhängigkeiten zwischen den Abtastwerten zu, so zeigt sich, daß die eindimensionale Verteilung am Eingang eines linearen Filters die Verteilung am Ausgang noch gar nicht determiniert, das heißt, bei gleicher Eingangsverteilung sind beliebig viele Ausgangsverteilungen möglich.

Der Grund für dieses Verhalten liegt in der Struktur linearer Systeme: Ein Filter mit einer Impulsantwort der Länge k bildet das Ausgangssignal durch eine lineare Kombination aus k vorangegangenen Abtastwerten.

Ist das Eingangssignal zufällig, so kann dies als Abbildung eines k -dimensionalen Ereignisraums in einen eindimensionalen aufgefaßt werden. Zur Berechnung der eindimensionalen Verteilung am Ausgang ist deshalb die Kenntnis der k -dimensionalen Eingangsverteilung nötig, wobei k die Länge der Impulsantwort ist. Bei nichtrekursiven (IIR-) Filtern muß daher die unendlich-dimensionale Eingangsverteilung bekannt sein.

In der vorliegenden Arbeit wurde die Fouriertransformierte der Wahrscheinlichkeitsdichte, die sogenannte charakteristische Funktion, benutzt.

Ist die multidimensionale charakteristische Funktion am Eingang des Filters bekannt, so erhält man eine geschlossene Lösung für die eindimensionale charakteristische Funktion am Ausgang mit Hilfe einer Variablentransformation. Die Wahrscheinlichkeitsdichte am Ausgang kann anschließend auf numerischem Wege mittels FFT gewonnen werden.

Für Filter mit unendlich langer Impulsantwort muß allerdings berücksichtigt werden, daß die charakteristische Funktion nur für endlich viele Dimensionen definiert ist.

Ein Grenzübergang führt hier zu der gesuchten Lösung. Das Problem der Berechnung der Verteilung nach der Filterung ist also mit Hilfe der vorgeschlagenen Methode immer dann lösbar, wenn die multidimensionale

charakteristische Funktion des Eingangsprozesses bekannt ist.

Die Bestimmung der multidimensionalen charakteristischen Funktion dürfte deshalb für die meisten praktischen Anwendungen das Hauptproblem darstellen.

Zur Demonstration des Verfahrens wurde die digitale Version des bereits beschriebenen quadratischen Detektors untersucht.

Liegt am Eingang eines solchen Detektors weißes gaußsches Rauschen an, so erhält man nach dem ersten Filter einen multidimensionalen Gaußprozeß, der durch die Impulsantwort des Filters beschrieben werden kann.

Nach der Quadratur haben wir dann eine multidimensionale, nichtzentrale χ^2 - Verteilung, deren charakteristische Funktion bekannt ist. Die Berechnung der Ausgangsverteilung ist nun für beliebige Ausgangsfilter möglich. Die Richtigkeit der Methode wurde mit Hilfe einer Monte-Carlo-Simulation untermauert.

Literaturverzeichnis

- [1] Anderson, T.W.: Introduction to multivariate statistical analysis, John Wiley & Sons, Inc., New York, London, Sydney, Toronto, 1958
- [2] Azizi, S.A.: Entwurf und Realisierung digitaler Filter, R. Oldenbourg Verlag, München-Wien-Oldenbourg, 1981.
- [3] Bergland, G.D.: A Fast Fourier Transform algorithm using base eight iterations, Math. Computation, vol.22, S.275-279, 1968
- [4] Bergland, G.D.: A radix-eight Fast Fourier Transform subroutine for real-valued series, IEEE Trans. Audio and Electroacoustics, vol.AU-17, Nr.2, S.138-144, 1969
- [5] Bilinsky, I.Ya., Vystavkin, A.N., Mikelson, A.K.: Processing of randomly-sampled signals, in: Young, I.T. et al. (Hrsg.) Signal Processing III, Band 1, S. 109-112, North Holland, Amsterdam-New York-Oxford-Tokyo, 1986
- [6] Box, G.E.P., Jenkins, G.M.: Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco, 1974.
- [7] Box, G.E.P., Muller, M.E.: A note on the generation of random normal deviates, Annals. Math. Stat., 29, 1958, S.610-611
- [8] Brenner, N.M.: Three FORTRAN Programs that perform the Cooley - Tukey Fourier Transform, Massachusetts Institute of Technology, Lincoln Laboratory, Lexington Massachusetts, Technical Note 1967-2.
- [9] Blumenson, L.E., Miller, K.S.: Properties of generalized Rayleigh distributions, Ann. Math. Statist., vol.34, S.903-910, 1963.

- [10] Brigham, E.O.: FFT: Schnelle Fourier Transformation, Übers. von S.A. Azizi, R. Oldenbourg Verlag, München-Wien-Oldenbourg, 1987, 3.Auflage.
- [11] Cooley, J.W., Tukey, J.W.: An algorithm for the machine calculation of complex Fourier series, Math. Computation, vol.19, S.297-301, 1965.
- [12] Courant, R., Hilbert, D.: Mathematische Methoden der Physik I, Springer Verlag, Berlin - Heidelberg - New York, 1968, 3. Auflage.
- [13] Cramér, H., Wold, H.: Some theorems on distribution functions, J. London Math. Soc., 11, S.290-294, 1936.
- [14] Cramér, H.: Mathematical methods of statistics, Princeton University Press, Princeton, N.J, 1946.
- [15] Davenport, B.D., Root, W.L.: An introduction to the theory of random noise, McGraw-Hill, New York, 1958.
- [16] Deutsch, R.: Nonlinear transformations of random processes, Prentice - Hall, Inc., Englewood Cliffs, New Jersey, 1962
- [17] Dittloff, H.-J.: Feldtheoretisch exakte Analyse und Optimierung von Mikrowellen-Multiplexern in integrierter Hohlleitertechnik, VDI Verlag, Reihe 21: Elektrotechnik, Nr.32, 1988
- [18] Doob, J.L.: Sochastic Processes, John Wiley & Sons, New York, London, Sydney, 1953.
- [19] Dudewicz E.J., Ralley, T.G.: The handbook of random number generation and testing with TESTRAND computer code, American Sciences Press, Inc, Columbus Ohio, 1981.

- [20] Dudgeon, D.E., Mersereau, R.M.: Multidimensional digital signal processing, Prentice – Hall, Inc., Englewood Cliffs, New Jersey, 1984
- [21] Duhamel, P., Hollmann, H.: Split Radix FFT algorithm, Electron. Letters, vol.20, S.14–16, 1984.
- [22] Duhamel, P.: Implementation of split-radix FFT algorithms for complex, real, and real-symmetric data, IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-34, S.285–295, 1986.
- [23] Eichenauer, J., Lehn, J.: A non-linear congruential pseudo random number generator, Statistical Papers, 28, 1986, S.315–326 Darmstadt,
- [24] Eichenauer, J., Grothe, H., Lehn, J., Topuzoglu, A.: A multiple rekursive non-linear congruential pseudo random number generator, Technische Hochschule Darmstadt, Fachbereich Mathematik, Preprint Nr. 1033, 1987
- [25] Emerson, R.C.: First probability densities for receivers with square law detectors, Journal of Applied Physics, vol.24, S.1168–1176, 1953
- [26] Fisz, M.: Wahrscheinlichkeitsrechnung und mathematische Statistik, VEB Deutscher Verlag der Wissenschaften, Berlin 1980, 10. Auflage
- [27] Gentleman, W.M., Sande, G.: Fast Fourier Transform – for fun and profit, Fall Joint Computer Conference, AFIPS Proc. vol.29, S. 563–578, 1966.
- [28] Gnedenko, B.: The theory of probability, Mir Publishers, Moskau, 1976.
- [29] Grube, A.: Mehrfach rekursiv erzeugte Zufallszahlen, Dissertation, Fakultät für Mathematik, Universität Karlsruhe, 1973

- [30] Halton, J.H.: A retrospective and prospective survey of the Monte Carlo method, SIAM Rev. Vol 12, 1-63, 1970.
- [31] Hamming, R. W.: Digital Filters, Prentice - Hall, Inc., Englewood Cliffs, New Jersey, 1977
- [32] Jensen, D.R.: Limit properties of noncentral multivariate Rayleigh and Chi-square Distributions, SIAM J. Appl. Math., vol.17, Nr.4, S.802-814, 1969.
- [33] Johnson N.L., Kotz S.: Distributions in Statistics: Continuous multivariate Distributions, Band 4 John Wiley & Sons, Inc., New York, London, Sydney, Toronto, 1972
- [34] Kac, M., Siebert, A.J.F.: On the theory of noise in radio receivers with square law detectors, Journal of Applied Physics, vol.18, S.383-397, 1947
- [35] Kaiser, J.F.: Digital filters, in: Kuo, F.F., Kaiser, J.F.: System analysis by digital computers, Kapitel 7, John Wiley & Sons, New York, 1966
- [36] Kendall, M., Stuart, A.S.: The advanced theory of statistics, Vol.1, Charles Griffins & Company Limited, London, 1970, 4.Auflage
- [37] Krishnamoorthy, A.S., Parthasarathy, M.: A multivariate Gamma Type distribution, Ann. Math. Statist., vol 22, S. 549-557, 1951
- [38] Krishnamoorthy, A.S., Parthasarathy, M.: Correction to "A multivariate Gamma Type distribution" , Ann. Math. Statist., vol 31, S.229, 1960

- [39] Knuth, D.E.: The art of computer programming, vol. 2, Addison-Wesley Publishing Company, Inc., Reading, Mass., 1981, 2. Auflage.
- [40] Lewis, T.G., Payne, W.H.: Generalized feedback shift register pseudo random number algorithm, Journal of the ACM, vol.20, Nr.3, S.456-468, 1973
- [41] Lukacs, E., Laha, R.G.: Applications of characteristic functions, Charles Griffins & Company Limited, London, 1964
- [42] Lukacs, E.: Characteristic functions, Charles Griffins & Company Limited, London, 1970, 2. Auflage
- [43] Mantei, A., Schreiber, C.: Amplitudendichteverteilung von Rauschen nach der Filterung, Diplomarbeit, Universität Bremen 1986 (unveröffentlicht).
- [44] Mantei, A., Schreiber, C., Simmer, K.U., Spöring, K: Probability density of band-limited noise after filtering, in: Young, I.T. et al. (Hrsg.) Signal Processing III, Band 1, S. 85-88, North Holland, Amsterdam-New York-Oxford-Tokyo, 1986
- [45] Marsaglia, G.: Random numbers fall mainly in the planes, Proc. Nat. Acad. Sci., Vol. 61, S. 25-28, 1968
- [46] Marsaglia, G.: The structure of linear congruential sequences, in: Zarembka, S.K. (Hrsg): Applications of number theory to numerical analysis, Academic Press, New York, London, 1972, S. 249-285
- [47] Metropolis, N., Ulam, S.: On the Monte Carlo method, J. Amer. Statist. Assoc. vol.44, S.25-44, 1949

- [48] Middleton, D.: An Introduction to statistical communication theory, McGraw-Hill, New York-Toronto-London 1960.
- [49] Meyer, M.A. Middleton, D., On the distribution of signals and noise after rectification and filtering, Journal of Applied Physics, vol. 25, S.1037-1052, 1954.
- [50] Nussbaumer, N.J.: Fast Fourier transform and convolution algorithms, Springer Verlag, Berlin - Heidelberg - New York, 1982, 2. Auflage,
- [51] Papoulis, A.: Probability, random variables and stochastic processes, McGraw-Hill, New York-Toronto-London 1965.
- [52] Roe, G.M., White, G.M.: Probability density for correlators with noisy reference Signals", I.R.E. Trans. PGIT, vol.IT-7, Nr.1, S.13-18, 1961
- [53] Schmitz, N., Lehmann, F.: Monte-Carlo-Methoden I, Verlag Anton Hain, Meisenheim am Glan, 1976
- [54] Schlitt, H: Systemtheorie für regellose Vorgänge, Springer Verlag, Berlin - Heidelberg - New York,, 1960
- [55] Schüßler, H.W.: Digitale Systeme zur Signalverarbeitung, Springer Verlag, Berlin - Heidelberg - New York, 1973
- [56] Schüßler, H.W.: Netzwerke, Signale und Systeme, Band II, Springer Verlag, Berlin - Heidelberg - New York, 1984
- [57] Siegert, A.J.F.: Passage of stationary Processes through linear and non-linear devices, I.R.E. Trans. PGIT, vol.3, Nr.1, S.4-25, 1954

- [58] Simmer, K.U., Spöring, K.: Berechnung der charakteristischen Funktion von gefiltertem Rauschen, Diplomarbeit, Universität Bremen 1986 (unveröffentlicht).
- [59] Sorensen, H.V., Heideman, M.T., Burrus, C.S.:
On computing the Split-Radix-FFT, IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-34, S.152-156, 1986.
- [60] Sorensen, H.V., Jones, D.L., Heideman, M.T., Burrus, C.S.:
Real-valued Fast Fourier Transform algorithms, IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-35, Nr.6, S.849-863, 1987.
- [61] Sorensen, H.V., Jones, D.L., Heideman, M.T., Burrus, C.S.:
Corrections to "Real-valued Fast Fourier Transform algorithms", IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-35, Nr.9, S. 1353, 1987.
- [62] System three hundred and sixty scientific subroutine package, Version 3. Programmers Manual. Program nr 360A-CM-03X, International Business Machines Corporation. 5. ed., New York 1970.
- [63] Uhlenbeck, G.E., Ornstein, L.S.: The Theory of the Brownian Motion. I, Phys. Rev., vol. 36, S.823-841, 1930.
- [64] Wang, M.C., Uhlenbeck, G.E.: On the Theory of the Brownian Motion. II, Revs. Modern Phys., vol.17, S.323-342, 1945.
- [64] Whittlesey, J.R.B.: On the multidimensional uniformity of pseudorandom generators, Comm. Assoc. Comp. Mach. Vol.12, Nr. 5, S.247, 1967.
- [66] Zurmühl, R.: Praktische Mathematik für Ingenieure und Physiker, Springer Verlag, Berlin - Heidelberg - New York, 1965, 5. Auflage.

